ELECTRICAL ENGINEERING TEXTS

PRINCIPLES OF RADIO ENGINEERING

ELECTRICAL ENGINEERING TEXTS

HARRY E. CLIFFORD, Consulting Editor

BERG · Heaviside's Operational Calculus

CHAFFEE · Theory of Thermionic Vacuum Tubes

COBINE · Gaseous Conductions

DAWES · Course in Electrical Engineering
Vol. I.—Direct Currents
Vol. II.—Alternating Currents
Industrial Electricity—Part I
Industrial Electricity—Part II

GLASGOW · Principles of Radio Engineering

LANGSDORF · Principles of Direct-current Machines
Theory of Alternating-current Machinery

LAWRENCE · Principles of Alternating Currents
Principles of Alternating-current Machinery

LAWS · Electrical Measurements

LYON · Applications of the Method of Symmetrical Components

MOON · The Scientific Basis of Illuminating Engineering

STEPHENS · The Elementary Theory of Operational Mathematics

ELECTRICAL ENGINEERING TEXTS

PRINCIPLES
OF
RADIO ENGINEERING

BY

R. S. GLASGOW, M.S.
Dean, U. S. Naval Postgraduate School

FIRST EDITION
NINETEENTH IMPRESSION

McGRAW-HILL BOOK COMPANY, INC.
NEW YORK AND LONDON
1936

PRINTED IN THE UNITED STATES OF AMERICA

THE MAPLE PRESS COMPANY, YORK, PA.

PREFACE

This book is intended primarily as a text for students of electrical engineering and is based upon the lecture notes used by the author for a number of years in a course on radio communication given at Washington University. The aim is to give a thorough presentation of the fundamentals of radio communication and the application of these principles to the art.

A knowledge of the fundamental laws of electricity and magnetism is assumed. Alternating-current theory is covered in Chapter I, and may be omitted if this material has already been covered in preceding courses. Mathematical developments are freely used throughout the book, but none are introduced that are not essential for an understanding of the principles involved. These developments do not extend beyond the mathematical preparation contained in the usual undergraduate engineering curriculum. Thus, the material on electromagnetic radiation in Chapter XIV has avoided the vector notation in favor of the less concise, but more generally understood, methods of integral calculus. A comprehensive list of problems is included at the end of each chapter. These, so far as possible, are based upon practical data in order to acquaint the student with the values of the quantities dealt with, in addition to testing his understanding of the principles involved.

There has been included considerable material on iron-core inductances, coupled circuits, graphical methods of determining amplifier performance, push-pull circuits, antennas, and radio-frequency transmission lines which is often dismissed with little more than passing mention. New graphical methods that have been developed by the author to determine the distortion present in amplifier circuits are also described. Numerous references to published papers, particularly to recent developments, are given

in the footnotes, in order to enable the student to secure more detailed information on subjects that are not completely described in the text.

The author wishes to express his appreciation of the helpful criticisms and suggestions of Professor H. E. Clifford, Dean of the Harvard Engineering School.

R. S. GLASGOW.

ST. LOUIS, MO.,
May, 1936.

CONTENTS

PAGE

PREFACE . . . v

CHAPTER I

ALTERNATING CURRENTS . . . 1
1. Introduction . . . 1
2. Sine Wave of Voltage . . . 1
3. Frequency; Radio and Audio . . . 2
4. Average and Effective Values . . . 3
5. Circuits Containing Resistance and Inductance . . . 4
6. Circuits Containing Resistance and Capacitance . . . 6
7. Resistance, Inductance, and Capacitance in Series . . . 8
8. Parallel Circuits . . . 9
9. Complex Quantities . . . 10
10. Addition and Subtraction of Vectors . . . 11
11. Representation of Impedances by Complex Numbers . . . 12
12. Multiplication and Division of Vectors . . . 12
13. Impedances in Parallel . . . 13
14. Example of the Solution of a Circuit Using Complex Quantities 14
15. Polar Notation of Vectors . . . 16
16. Alternating-current Power . . . 19
17. Nonsinusoidal Waves . . . 21

CHAPTER II

RESONANT CIRCUITS . . . 28
18. Series Resonance . . . 28
19. Sharpness of Resonance . . . 29
20. Parallel Resonance . . . 32
21. Parallel Resonance When the Coil Contains Resistance . . . 35
22. Parallel Resonance When Both Coil and Condenser Contain Resistance . . . 44

CHAPTER III

PROPERTIES OF COILS AND CONDENSERS . . . 48
23. Inductance of a Single-layer Solenoid . . . 48
24. Multilayer Coils. Distributed Capacitance . . . 49
25. Coil Resistance. Skin Effect . . . 53
26. Properties of Iron-core Coils . . . 55
27. Magnetic Alloys . . . 60
28. Incremental Permeability . . . 65
29. Inductance of Iron-core Coils with Direct and Alternating Current Superimposed . . . 67

PAGE
30. Condensers 69
31. Variable Condensers 71
32. Electrolytic Condensers 73

CHAPTER IV

COUPLED CIRCUITS 77
33. Mutual Inductance 77
34. Transformers 79
35. Coupled Resonant Circuits 80
36. Antenna-circuit Adjustments in Receiving Sets 90
37. Other Forms of Coupling 92

CHAPTER V

OSCILLATORY CIRCUITS 95
38. Free Oscillations and Mechanical Analogies 95
39. Charge of a Condenser through an Inductance and Resistance in Series 95
40. Discharge of a Condenser through an Inductance and Resistance 102
41. Free Oscillations in Coupled Circuits 103
42. Spark Transmitters 106
43. Radio Interference from Electrical Sparks 110
44. Limitations of Spark Transmitters 111

CHAPTER VI

FUNDAMENTAL PROPERTIES OF VACUUM TUBES 115
45. Introduction 115
46. History of the Thermionic Tube 115
47. Emission of Electrons 117
48. Types of Cathodes 121
49. Comparison of Emitters 125
50. Effects of Gas 127
51. Triodes 129
52. Characteristic Curves of Triodes 130
53. Triode Constants 131
54. Expressions for the Plate Current 134
55. Vacuum-tube Notation 136
56. Equivalent Circuit of a Triode 137
57. Calculation of Triode Constants from the Structural Dimensions 139
58. Measurement of Triode Constants 140

CHAPTER VII

AUDIO-FREQUENCY AMPLIFIERS 144
59. Vacuum-tube Amplifiers and Their Classification 144
60. Distortion in Amplifiers 146
61. Amplifier with Resistance Load 150
62. Resistance-coupled Amplifiers 154
63. Impedance-coupled Amplifiers 157

PAGE
64. Amplifier with an Inductive Load 161
65. Transformer-coupled Amplifier. 163
66. Transformer Characteristics. 166
67. Comparison of Amplifiers 169
68. Amplification Expressed in Decibels 170
69. Measurement of Voltage Amplification 171
70. Considerations for Maximum Power 172
71. Power Amplifiers. 177
72. Graphical Determination of Power Output and Distortion . . 179
73. Approximate Determination of the Output of a Triode 183
74. Push-pull Amplifiers 184
75. Advantages of Push-pull Amplifiers. 190
76. Methods of Obtaining C bias 192
77. Push-pull Amplifiers, Class B 195
78. Push-pull Amplifiers, Class AB. 200
79. Amplifiers Using Tetrodes. 201
80. Tetrodes with Space-charge Grid. 204
81. Power Amplifiers Using Pentodes. 204
82. Graphical Determination of Distortion and Power Output of Pentodes 207
83. Direct-coupled Amplifiers. 209
84. Feed-back in Amplifiers. 212
85. Prevention of Feed-back 215
86. Amplifiers Employing Feed-back. 217
87. Telephone Repeaters 220

CHAPTER VIII

INPUT IMPEDANCE OF A TRIODE. 228
88. Equivalent Circuit of a Triode. 228
89. Input Impedance of the Grid Circuit. 228
90. Case 1. $Z_b = R_b$. 230
91. Case 2. $Z_b = R_b + j\omega L_b$ 232
92. Case 3. $Z_b = R_b - jX_b$. 235

CHAPTER IX

RADIO-FREQUENCY AMPLIFIERS FOR RECEPTION. 237
93. Types of Amplifiers. 237
94. Untuned Amplifiers. 237
95. Tuned Amplifiers. 239
96. Effect of Mutual Inductance. 243
97. Combinations of Inductive and Capacitive Coupling 246
98. Amplifiers Using Tuned Coupled Circuits 249
99. Cascade Amplifiers. 251
100. Methods of Avoiding Oscillation. 253
101. Neutralizing Circuits. 254
102. Neutralizing Adjustments. 258

PAGE

CHAPTER X

OSCILLATORS AND RADIO-FREQUENCY POWER AMPLIFIERS 261
103. Oscillators 261
104. Vacuum-tube Oscillators 261
105. Oscillator Circuits 266
106. Current and Voltage Relations 269
107. Circuit Calculations 273
108. Power Relations 278
109. Power-amplifier Computations Based upon Approximate Tube Characteristic 283
110. Amplifier-tuning Adjustments 285
111. Class B Amplifiers 287
112. Class C Amplifiers 289
113. Frequency Multipliers 291
114. Neutralization of Power Amplifiers 293
115. Crystal Oscillators 294
116. Magnetostriction Oscillators 299
117. Frequency Stability of Oscillators 299
118. Beat-frequency Oscillators 301
119. The Multivibrator 303
120. Dynatron Oscillators 304
121. Oscillating Arc 306
122. Magnetron Oscillators 307
123. Barkhausen Oscillations 309

CHAPTER XI

MODULATION 313
124. Types of Modulation 313
125. Amplitude Modulation 315
126. Energy Relations 317
127. Methods of Obtaining Amplitude Modulation 318
128. Absorption Modulation 318
129. Plate-modulated Oscillators 319
130. Plate-modulated Amplifiers 321
131. Modulator Coupling Circuits 323
132. Grid-modulated Amplifiers 326
133. Modulation Due to a Nonlinear Impedance 327
134. The van der Bijl Type of Modulated Amplifier 330
135. Balanced Modulators for Carrier Suppression 331
136. Grid-current Modulation 333
137. Single-side-band Transmission 334
138. Other Types of Modulators 336
139. Plate Current Expressed as an Infinite Series 337

CHAPTER XII

VACUUM-TUBE DETECTORS 344
140. Operation of Detectors 344
141. Plate Detection 345

PAGE
142. Grid-current Detection 349
143 Frequency Distortion in Grid-leak Detectors 355
144. Grid-leak Power Detectors 358
145. Grid-bias Detectors 360
146. Diode Detectors 365
147. Heterodyne Detectors 370
148. Regenerative Detectors 372
149. Superregeneration 374
150. Vacuum-tube Voltmeters 375

CHAPTER XIII

RECEIVING SYSTEMS 380
151. Types of Receivers 380
152. Tuned Radio-frequency Receivers 380
153. The Superheterodyne 381
154. Image-frequency Interference 384
155. Choice of Intermediate Frequency 385
156. Frequency Converters 386
157. Electron-coupled Frequency Converters 388
158. Variable-mu Tubes 390
159. Automatic Volume Control 394
160. Interchannel-noise Suppression 398
161. The Compandor 400
162. Acoustically Compensated Volume Control 402
163. Tone Control 403
164. Typical Broadcast-receiving Sets 404
165. Performance Tests on Receiving Sets 407
166. Loud-speakers 410

CHAPTER XIV

ANTENNAS AND WAVE PROPAGATION 413
167. Electromagnetic Waves 413
168. Electrical Units 415
169. Current and Voltage Distribution in an Antenna 416
170. Electrical Images 418
171. Electromagnetic Radiation 419
172. Field Distribution of Vertical Quarter-wave Antenna 420
173. Field Distribution around Vertical Antenna 424
174. Radiation Resistance 429
175. Wave Length for Optimum Ground Wave 433
176. Directional Antennas 435
177. Loop Antennas 436
178. Radio Direction Finders 439
179. Radio-range Beacons 443
180. Principles of Antenna Arrays 449
181. Determination of the Polar Diagram for an Array 454
182. Types of Arrays 458
183. Arrays for Reception 460

PAGE

184. Practical Aspects of Antenna Arrays 461
185. Directional Antennas Employing Long Wires 463
186. The Beverage or Wave Antenna 465
187. Tilted-wire and Inverted-V Antennas. 468
188. Horizontal Rhombic or Diamond-shaped Antenna 471
189. Multiple-tuned Antenna. 473
190. Radio-frequency Transmission Lines 474
191. Resonant Transmission Lines 475
192. Nonresonant Transmission Lines. 477
193. Transmission-line Theory 481
194. Characteristics of Lines Having Uniformly Distributed Constants. 484
195. Quarter-wave-length Line as an Impedance-matching Device . 486
196. Transmission-line Measurements. 489
197. Transmission Lines for Reception 489
198. Propagation of Electromagnetic Waves 491
199. The Ionosphere 492
200. Skip-distance Effect 495
201. Echoes 496
202. Fading 497
203. Long-wave Transmission 499
204. Critical Wave Lengths 499
205. Short-wave Transmission 500

NAME INDEX. 505

SUBJECT INDEX 509

PRINCIPLES OF RADIO ENGINEERING

CHAPTER I

ALTERNATING CURRENTS

1. Introduction.—Radio communication is primarily based upon alternating-current theory. It is the purpose of this chapter to present in condensed form the laws and definitions which are directly applicable to the various types of circuits employed in radio apparatus and to the numerous by-products of the art. For a more extended treatment of these principles the student is referred to the numerous texts in electrical engineering in which the subject is treated in greater detail.

2. Sine Wave of Voltage.—If a coil of N turns is rotated with a constant angular velocity of ω radians per second in a uniform magnetic field, the electromotive force (e.m.f.) induced in it at any instant will be

$$e = -N\frac{d\phi}{dt} \text{ abvolts} \tag{1}$$

The magnetic flux ϕ passing through the coil will be a maximum when the plane of the coil is perpendicular to the flux. When the coil has rotated to a position parallel to the magnetic field, the flux linking with, or passing through, the coil will be zero. Obviously, the flux linking with the coil for any intermediate position will vary according to a cosine law and may be expressed by

$$\phi = \Phi_m \cos \omega t \tag{2}$$

where t is the time in seconds required for the coil to move from the perpendicular position to the location in question. Differentiating (2) and substituting in (1), we obtain

$$e = \omega N \Phi_m \sin \omega t \tag{3}$$

The maximum value of e will occur when $\sin \omega t$ is unity or when $\omega t = \pi/2$, so that

$$E_m = \omega N \Phi_m \text{ abvolts} = \omega N \Phi_m \times 10^{-8} \text{ volts} \tag{4}$$

Substituting (4) in (3),

$$e = E_m \sin \omega t \tag{5}$$

Most commercial alternators produce voltage waves which are approximately sinusoidal. Occasionally the wave form deviates considerably from a sine wave, as shown in Fig. 1. The treatment of distorted wave shapes will be discussed later.

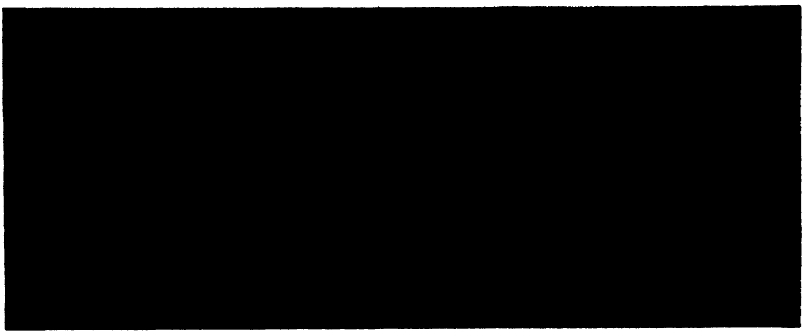


FIG. 1.—Oscillogram of distorted wave.

3. Frequency; Radio and Audio.—The number of complete cycles through which an alternating current or e.m.f. passes in 1 sec. is called its *frequency*. Thus, the coil in Sec. 2 will produce one cycle for each complete revolution, so that

$$f = \frac{\omega}{2\pi} \tag{6}$$

The range of frequencies employed in radio apparatus extends from a few cycles per second to values as high as 10^9 cycles. It is more convenient to express the higher frequencies in kilocycles, one kilocycle being one thousand cycles. The lower frequencies up to about 10 kc are commonly designated as *audio frequencies*. Currents in this frequency range would produce audible sound if passed through a suitable telephone receiver. *Radio frequencies* are those which lie above the audio-frequency band. The average ear becomes deaf to frequencies above 15 kc so that

there is some overlap between audio and radio frequencies in the lower range of the latter.

4. Average and Effective Values.—The average value of an alternating current for a complete cycle is zero. Thus, the ordinary direct-current instrument, which reads average values, will read zero when connected in an alternating-current circuit. The usual meaning of average value as applied to alternating currents is the average ordinate of a half wave. This would be the area of the wave divided by its base, so that if the current at any instant is expressed by $i = I_m \sin \omega t$, the average value of this current will be

$$\begin{aligned} I_{\text{avg}} &= \frac{1}{\pi}\int_0^\pi I_m \sin \omega t \cdot d\omega t \\ &= \frac{2I_m}{\pi} = 0.637 I_m \end{aligned} \tag{7}$$

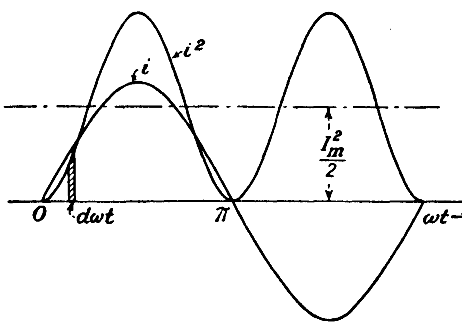


FIG. 2.—Current and current-squared waves.

The average value of an alternating current or voltage is of minor importance, the *effective* or *root-mean-square* (r.m.s.) value being the one with which we are chiefly concerned. The *effective value of an alternating current is that value of direct current which will produce the same heating effect.* When an alternating current $i = I_m \sin \omega t$ flows through a resistance R, heat is dissipated in the resistance which is proportional to i^2R. The average rate of transformation of electrical energy into heat will be proportional to the area under the curve of i^2 in Fig. 2 divided by the base, and is given by

$$I_{\text{eff}}^2 R = \frac{1}{\pi}\int_0^\pi i^2 R\, d\omega t = \frac{1}{\pi}\int_0^\pi I_m^2 R \sin^2 \omega t \cdot d\omega t \tag{8}$$

The effective value of the current will be

$$I_{\text{eff}} = \sqrt{\frac{I_m^2}{\pi}\int_0^{\pi} \sin^2 \omega t \cdot d\omega t} = \frac{I_m}{\sqrt{2}} = 0.707 I_m \tag{9}$$

In a similar manner

$$E_{\text{eff}} = \frac{E_m}{\sqrt{2}} \tag{10}$$

From the definition of effective value it follows that the average power consumed in an alternating-current circuit is given by

$$P = I_{\text{eff}}^2 R \tag{11}$$

where R is the resistance through which the current is flowing.

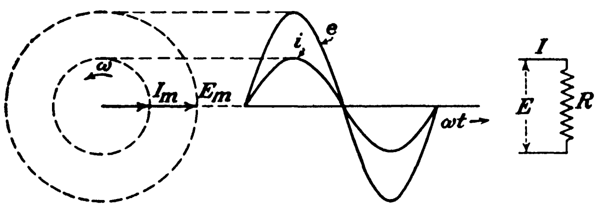


FIG. 3.—Current and voltage relations in a circuit containing pure resistance.

5. Circuits Containing Resistance and Inductance.—If an alternating current $i = I_m \sin \omega t$ flows through a pure resistance R, the voltage across the resistance will be

$$e = iR = I_m R \sin \omega t = E_m \sin \omega t \tag{12}$$

The current and voltage will be in phase with each other as shown in Fig. 3. These waves may be thought of as having been generated by the two vectors rotating counterclockwise at an angular velocity of ω. The lengths of these vectors are equal to the maximum values of the current and voltage.

If an alternating voltage is impressed across a resistance of R ohms in series with an inductance of L henrys, then from Kirchhoff's laws we have

$$e = iR + L\frac{di}{dt} \tag{13}$$

If the current through the circuit is given by $i = I_m \sin \omega t$, (13) becomes

$$e = I_m R \sin \omega t + I_m \omega L \cos \omega t \tag{14}$$

That is, the impressed voltage may be thought of as being made up of two terms, as shown in Fig. 4; the first being in phase with the current and the other being 90 degrees in advance of the current. The impressed voltage at any instant will be the instantaneous sum of these two components and will be ahead of the current by the angle θ. The equation for the voltage will be

$$e = E_m \sin (\omega t + \theta) \tag{15}$$

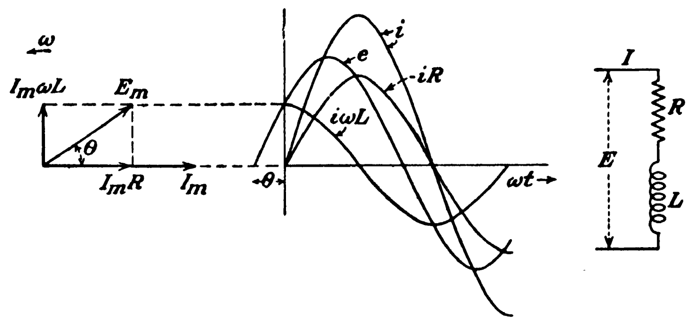


FIG. 4.—Current and voltage relations in a circuit containing resistance and inductance.

The vector diagram of the current and voltage relations is shown at the left in Fig. 4. These vectors illustrate the geometrical relations between the various quantities much more clearly than does the group of sine waves. Consequently, vector diagrams are usually used to show magnitude and phase relations of alternating-current quantities. The lengths of these vectors should be equal to the maximum values of the quantities they represent, if they are to be thought of as generating the waves of current and voltage by their rotation. However, effective values may be used instead, since the geometry of the diagram is in no way altered.

The angle by which the current lags behind the impressed voltage is

$$\theta = \tan^{-1} \frac{\omega L}{R} \tag{16}$$

In the event that the resistance of the circuit was zero, the angle of lag would be 90 degrees.

The factor ωL is termed the *inductive reactance* and is denoted by X_L, so that

$$X_L = \omega L = 2\pi f L \tag{17}$$

and is expressed in *ohms* if L is in henrys.

The *impedance* of an alternating-current circuit is defined as *the ratio of the applied voltage to the resultant steady-state current.* It is also expressed in ohms and is denoted by Z. Dividing the magnitudes in the vector diagram of Fig. 4 by I_m, the relations shown in Fig. 5 are obtained, so that

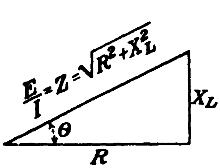


FIG. 5.—Relations between resistance, reactance, and impedance.

$$Z = \sqrt{R^2 + X_L^2} \tag{18}$$

Thus, a circuit composed of 8 ohms resistance in series with 6 ohms of inductive reactance would have an impedance of 10 ohms. The circuit constants are no longer added arithmetically as is the case in direct currents.

6. Circuits Containing Resistance and Capacitance.—When an alternating current of $i = I_m \sin \omega t$ flows through a resistance and capacitance in series, the voltage required to maintain this current at any instant will be

$$e = iR + \frac{q}{C} \tag{19}$$

where q is the charge expressed in coulombs and C is the capacitance in farads. The relation between charge and current is

$$q = \int i dt \tag{20}$$

Substituting the equation of the current and the above expression for q in (19), we obtain

$$\begin{aligned} e &= I_m R \sin \omega t + \frac{1}{C}\int I_m \sin \omega t dt \\ &= I_m R \sin \omega t - \frac{I_m}{\omega C} \cos \omega t \end{aligned} \tag{21}$$

It is seen that the impressed voltage is again composed of a sine and cosine term as in (14), but in this case the cosine term is negative. Figure 6 shows the phase relations of the various voltages and current. The effect of capacitance is to cause the current to lead the impressed voltage by the angle θ, which is directly opposite to the effect of inductance. If the resistance in the circuit were zero, the current would be 90 degrees ahead in phase of the impressed voltage.

The phase angle is now

$$\theta = \tan^{-1} \frac{1}{\omega CR} \tag{22}$$

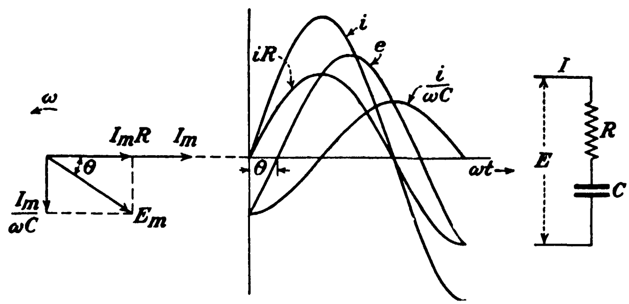


FIG. 6.—Current and voltage relations in a circuit containing resistance and capacitance.

The vector diagram is similar to the inductive case except that it is now inverted. The term $1/\omega C$ is the *capacitive* or *condensive reactance* and is denoted by X_C, so that

$$X_C = \frac{1}{\omega C} = \frac{1}{2\pi f C} \tag{23}$$

and is expressed in ohms if C is in farads.

The impedance of this circuit is

$$Z = \sqrt{R^2 + X_C^2} \tag{24}$$

which is identical in form with (18).

The farad is such an exceedingly large unit that in ordinary electrical work the capacitance of condensers is usually expressed in microfarads (μf), one microfarad being 10^{-6} farad. In radio communication the microfarad is often inconveniently large so

that the micro-microfarad ($\mu\mu$f) is frequently used, one micro-microfarad being 10^{-12} farad.

7. Resistance, Inductance, and Capacitance in Series.—Inductance in a circuit causes the current to lag behind the impressed voltage while capacitance results in a leading current. When both of these elements are connected in series, they tend to annul each other so that the resultant reactance of the circuit becomes the difference between X_L and X_C. The resultant impedance of the circuit is

$$Z = \sqrt{R^2 + (X_L - X_C)^2} \tag{25}$$

where R is the total circuit resistance, including any external nonreactive resistance as well as the individual resistances of the coil and condenser.

The vector diagram of the circuit is shown in Fig. 7. It will be observed that the voltage across the inductance (IX_L) is somewhat greater than the impressed voltage, E. If X_L and X_C are both large compared to R and are approximately equal in magnitude, the voltages across both coil and condenser can be greatly in excess of the applied voltage. It is, therefore, evident that the algebraic sum of the voltages across the various parts of the circuit is no longer equal to the impressed voltage, as is the case with direct currents. However, the *vector sum* of these voltages will be equal to the impressed voltage. The method of attack in alternating-current-circuit problems is exactly the same as in direct currents except that vector addition or subtraction must be used. Thus, if a number of impedances are connected in series the total impedance of the combination is

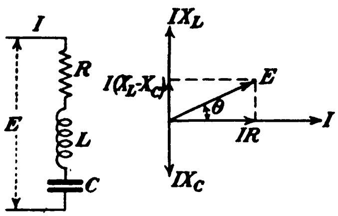


FIG. 7.—Vector diagram of a circuit containing R, L, and C in series.

$$\begin{aligned} Z_T &= \sqrt{(R_1+R_2 \cdots)^2+(X_{L_1}+X_{L_2} \cdots -X_{C_1}-X_{C_2} \cdots)^2} \\ &= \sqrt{R_T^2 + X_T^2} \end{aligned} \tag{26}$$

The phase angle of the circuit will be

$$\theta = \tan^{-1} \frac{X_T}{R_T} \tag{27}$$

The resultant current will either lead or lag the impressed voltage, depending on whether condensive or inductive reactance predominates. Should the total reactance be zero—in which case the circuit is said to be in *resonance*—the current will be in phase with the impressed voltage and will be limited only by the resistance in the circuit. Resonant circuits are of fundamental importance in radio and will be considered in detail in Chap. II.

8. Parallel Circuits.—When two impedances Z_1 and Z_2 are connected in parallel as in Fig. 8, the individual currents are

$$I_1 = \frac{E}{\sqrt{R_1^2 + X_1^2}}$$

$$I_2 = \frac{E}{\sqrt{R_2^2 + X_2^2}}$$

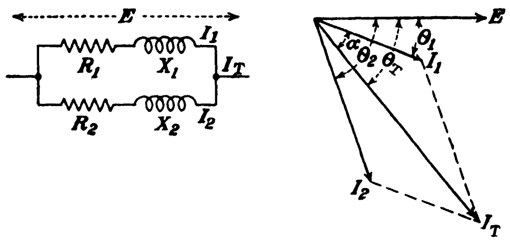


FIG. 8.—Vector diagram of a parallel circuit.

The total current will be the vector sum of I_1 and I_2 and may be obtained from

$$I_T = \sqrt{I_1^2 + I_2^2 + 2I_1I_2 \cos(\theta_2 - \theta_1)} \tag{28}$$

The phase angle between E and I_T is

$$\theta_T = \theta_1 + \alpha.$$

The angle α can be determined from the law of cosines by the relation

$$\cos \alpha = \frac{I_1^2 + I_T^2 - I_2^2}{2I_1I_T}.$$

If one of the impedances had possessed condensive reactance so that the current through it led the impressed voltage, the included angle between I_1 and I_2 would then have been $\theta_1 + \theta_2$. The sum of the two angles would therefore be used in (28) instead of their difference.

The above calculations are somewhat cumbersome and are not conveniently adapted to slide-rule computations. Furthermore, if a number of impedances are connected in parallel, the labor involved is considerable so that a more convenient method is required. The use of complex algebra in such problems results in a considerable simplification of the work involved.

9. Complex Quantities.—Vectors representing alternating currents and voltages may be referred to a system of coordinate axes and expressed in terms of their components along these axes.

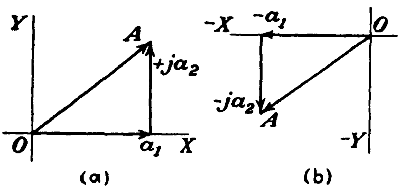


FIG. 9.—Representation of vectors by complex numbers.

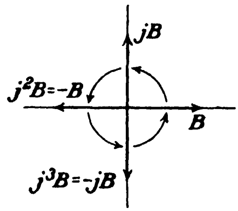


FIG. 10.—Rotation of a vector by the operator j.

Thus, in Fig. 9*a* the vector A has a component a_1 units along the axis OX and a_2 units along the axis OY. In order to distinguish between the components in the two directions, those along the OY axis will be prefixed by the symbol j. The vector A, in terms of its components, can be written

$$A = a_1 + ja_2 \tag{29}$$

where the addition must be considered in a vector sense. The scalar magnitude of the vector will be

$$|A| = \sqrt{a_1^2 + a_2^2} \tag{30}$$

the vertical bars being used to signify the absolute or scalar magnitude of the quantity they enclose.

Components of a vector to the left or below the origin O would have negative signs. Thus, a vector A drawn in the third quadrant as in Fig. 9*b* would be expressed by

$$A = -a_1 - ja_2 \tag{31}$$

Its length or scalar magnitude would be given by (30), as before.

The symbol j is an operator which rotates counterclockwise the quantity to which it is attached through an angle of 90 degrees. If the vector B in Fig. 10 is multiplied by j, it will then point vertically upward. Multiplying again by j should produce a further rotation of 90 degrees. But in this position the vector now extends to the left of the origin in a negative direction so that $j^2B = -B$, or $j^2 = -1$, so that

$$j = \sqrt{-1} \tag{32}$$

which is mathematically an imaginary number. For this reason the vertical component of a vector is frequently termed the "imaginary component." This expression is somewhat misleading and the expression "quadrature component" is preferable.

If the vector $j^2B = -B$ is again multiplied by the operator j, another 90 degrees rotation results so that it now points vertically downward. A fourth application of the operator j brings the vector back to its original position.

10. Addition and Subtraction of Vectors.—Vector addition is accomplished by merely adding the horizontal and quadrature components respectively, with due regard to the algebraic sign preceding the component. The algebraic sums of the horizontal and vertical components become the respective components of the resultant. If the two currents in Fig. 8 had been expressed in complex form, as in Fig. 11, the determination of the total current would have been greatly simplified.

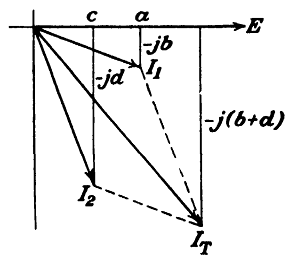


FIG. 11.—Addition of currents expressed in complex form.

Let the vector expressions for the two currents be

$$I_1 = a - jb$$
$$I_2 = c - jd$$

The vector expression for the total current will be

$$\dot{I}_T = \dot{I}_1 + \dot{I}_2 = (a + c) - j(b + d) \tag{33}$$

and its scalar magnitude is given by

$$|I_T| = \sqrt{(a + c)^2 + (b + d)^2} \tag{34}$$

The dots placed over the currents in (33) are frequently employed to indicate that the quantities are vectors and are to be combined vectorially. Since the majority of expressions encountered in alternating-current circuits are vector expressions, the use of such symbols is usually unnecessary. Where confusion is apt to result between scalars and vectors, the scalar quantity will be indicated by the vertical bars, as in (34).

The subtraction of a vector is accomplished by merely reversing its direction and then combining it as in addition. This is performed analytically by reversing the signs preceding the two components and then adding.

11. Representation of Impedances by Complex Numbers.—It has already been shown (Fig. 5) that the impedance of a circuit composed of resistance and inductive reactance may be represented by the hypotenuse of a right triangle whose base and altitude are R and X_L, respectively. This can be expressed in complex form by

$$Z = R + jX_L \tag{35}$$

Since condensive reactance is opposite to inductive reactance in its effects, the impedance of a circuit composed of resistance and condensive reactance would be represented by

$$Z = R - jX_C \tag{36}$$

It should be realized that impedances are not vector quantities and that (35) and (36) are merely methods of representing the geometrical relations involved. While neither resistance nor reactance is a vector, resistance and reactance drops (IR and IX) are vectors. Impedances, therefore, must be handled like vectors so far as the operations of addition, subtraction, multiplication, and division are concerned.

12. Multiplication and Division of Vectors.—Suppose a current $I = i_1 + ji_2$ is to be multiplied by an impedance $Z = r + jx$. Ordinary algebraic procedure is followed, bearing in mind that $j^2 = -1$. Thus,

$$\begin{aligned} IZ &= (i_1 + ji_2)(r + jx) = i_1r + ji_1x + ji_2r + j^2i_2x \\ &= (i_1r - i_2x) + j(i_1x + i_2r) \end{aligned} \tag{37}$$

If the reactance in the circuit had been condensive,

$$IZ = (i_1 + ji_2)(r - jx) = (i_1r + i_2x) + j(i_2r - i_1x) \tag{38}$$

The magnitude of the voltage across the impedance will be the square root of the sum of the squares of the two components in (37) or (38).

Division is accomplished as follows: Let it be required to find the vector expression for the current resulting from a voltage $E = e_1 + je_2$ being impressed across an impedance $Z = r + jx$. The current will be

$$I = \frac{E}{Z} = \frac{e_1 + je_2}{r + jx}$$

If we multiply numerator and denominator by the denominator with the sign before the j term reversed so as to rationalize the denominator, we have

$$\begin{aligned} I &= \frac{e_1 + je_2}{r + jx} \times \frac{r - jx}{r - jx} = \frac{e_1r - je_1x + je_2r - j^2e_2x}{r^2 - j^2x^2} \\ &= \frac{(e_1r + e_2x) + j(e_2r - e_1x)}{r^2 + x^2} = \frac{e_1r + e_2x}{r^2 + x^2} + j\frac{e_2r - e_1x}{r^2 + x^2} \end{aligned} \tag{39}$$

Had the impedance been $Z = r - jx$, the current would have been

$$I = \frac{e_1 + je_2}{r - jx} \times \frac{r + jx}{r + jx} = \frac{e_1r - e_2x}{r^2 + x^2} + j\frac{e_2r + e_1x}{r^2 + x^2} \tag{40}$$

In a simple series circuit where only the magnitude of the current or voltage is sought there is no particular advantage in using complex quantities. In the preceding example it would have been easier to have found the current by dividing the scalar magnitude of the voltage by the scalar magnitude of the impedance, which would have given the current in amperes. The only advantage of the complex notation is in parallel or series-parallel circuits.

13. Impedances in Parallel.—As in direct currents, where the resistance R of a number of resistances in parallel is given by

$$\frac{1}{R} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} + \cdots \tag{41}$$

the resultant impedance Z of a number of impedances in parallel is given by the vector expression

$$\frac{1}{Z} = \frac{1}{z_1} + \frac{1}{z_2} + \frac{1}{z_3} + \cdots \tag{42}$$

The reciprocal of an impedance is called an *admittance* and is designated by the letter Y. It is measured in *mhos*. Thus (42) can be written

$$Y = y_1 + y_2 + y_3 + \cdots \tag{43}$$

When only two impedances in parallel are involved, the resultant impedance of the combination from (42) becomes

$$Z = \frac{Z_1 Z_2}{Z_1 + Z_2} = R_a \pm jX_a \tag{44}$$

where R_a and X_a are the apparent resistance and apparent reactance, respectively, and are the real and quadrature (or imaginary) components of the expression preceding. In other words, the parallel circuit could be replaced by a single impedance having

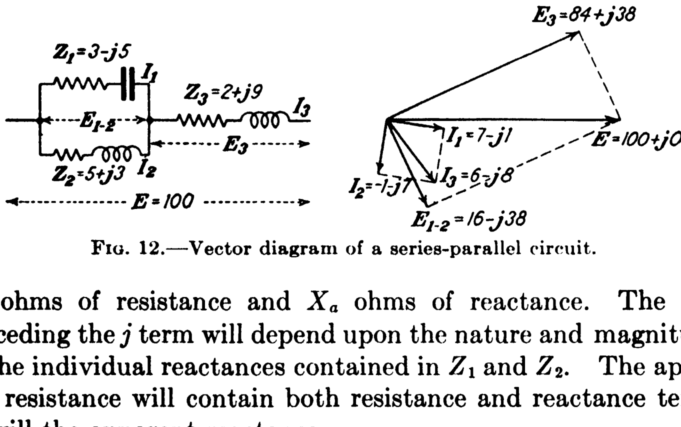


FIG. 12.—Vector diagram of a series-parallel circuit.

R_a ohms of resistance and X_a ohms of reactance. The sign preceding the j term will depend upon the nature and magnitudes of the individual reactances contained in Z_1 and Z_2. The apparent resistance will contain both resistance and reactance terms, as will the apparent reactance.

Complex expressions for the impedance must be used in (42), (43), and (44). The relations expressed by these equations no longer hold if the scalar magnitudes of the impedances are used.

14. Example of the Solution of a Circuit Using Complex Quantities.—An impedance Z_1 composed of 3 ohms resistance and 5 ohms condensive reactance is in parallel with a coil Z_2 having 5 ohms resistance and 3 ohms reactance. In series with this combination is an impedance Z_3 having 2 ohms resistance and 9 ohms inductive reactance. If 100 volts is impressed across the entire circuit, what will be the current in each branch? What

will be the voltage across each branch? The circuit diagram is shown in Fig. 12.

Since only the magnitude of the impressed voltage is given and none of the other currents or voltages are vectorially defined, we are at liberty to choose the vector representing the impressed voltage in any position we like. For convenience we shall make it the reference vector and its expression will then be

$$E = 100 + j0$$

The resultant impedance of Z_1 and Z_2 is

$$Z_{1-2} = \frac{Z_1 Z_2}{Z_1 + Z_2} = \frac{(3 - j5)(5 + j3)}{8 - j2} = \frac{(30 - j16)(8 + j2)}{8^2 + 2^2}$$
$$= \frac{272 - j68}{68} = 4 - j1 \text{ ohms.}$$

The total circuit impedance will be

$$Z_t = Z_{1-2} + Z_3 = (4 - j1) + (2 + j9) = 6 + j8 \text{ ohms.}$$

The total current is

$$I_3 = \frac{E}{Z_t} = \frac{100 + j0}{6 + j8} = \frac{100(6 - j8)}{6^2 + 8^2} = 6 - j8 \text{ amperes.}$$

The voltage drop across Z_3 is

$$E_3 = I_3 Z_3 = (6 - j8)(2 + j9) = 84 + j38 \text{ volts.}$$

The voltage drop across the parallel branch is

$$E_{1-2} = E - E_3 = (100 - j0) - (84 + j38) = 16 - j38 \text{ volts.}$$

The current through Z_1 is

$$I_1 = \frac{E_{1-2}}{Z_1} = \frac{16 - j38}{3 - j5} = \frac{(16 - j38)(3 + j5)}{3^2 + 5^2}$$
$$= \frac{238 - j34}{34} = 7 - j1 \text{ amperes.}$$

The current through I_2 may be obtained by

$$I_2 = I_3 - I_1 = (6 - j8) - (7 - j1) = -1 - j7 \text{ amperes.}$$

As a check on the solution, the current I_2 is also

$$I_2 = \frac{E_{1-2}}{Z_2} = \frac{16 - j38}{5 + j3} = \frac{(16 - j38)(5 - j3)}{5^2 + 3^2}$$
$$= \frac{-34 - j238}{34} = -1 - j7 \text{ amperes.}$$

The vector diagram is shown to the right of the circuit diagram in Fig. 12. The scalar magnitudes of any of the above currents and voltages can be obtained by taking the square root of the sum of the squares of the two components of the vector expression. Thus,

$$|I_3| = \sqrt{6^2 + 8^2} = 10 \text{ amperes.}$$

15. Polar Notation of Vectors.—Instead of defining a vector in terms of rectangular coordinates, polar coordinates may be used. The process of multiplication and division is greatly simplified with this form of notation, as will be seen later. In order that a vector may be completely defined, its magnitude and direction must be given. Thus, in Fig. 13*a* the vector has a length of A units and makes an angle θ with the reference axis. It may be written as

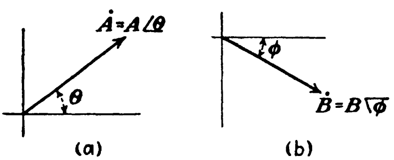


FIG. 13.—Polar representation of vectors.

$$\dot{A} = A\underline{/\theta} \tag{45}$$

Since in rectangular form the vector would be written

$$\dot{A} = a + jb \tag{46}$$

where

$$a = A \cos \theta \tag{47}$$
$$b = A \sin \theta \tag{48}$$
$$A = \sqrt{a^2 + b^2} \tag{49}$$

the transformation from polar to rectangular form is

$$\dot{A} = A(\cos \theta + j \sin \theta) \tag{50}$$

Likewise, a vector in the fourth quadrant as in Fig. 13*b* would be written

$$\dot{B} = B\overline{\diagdown\phi} = B\underline{/-\phi} \tag{51}$$
$$= B(\cos\phi - j\sin\phi) \tag{52}$$

Angles below the reference axis are to be regarded as negative and must be so treated when combined with positive angles. They may be considered as having been produced by negative (clockwise) rotation of the vector in question. Angles above the reference axis are treated as positive angles.

Multiplication of two polar vectors is performed by multiplying their magnitudes and adding their phase angles, with due regard to the sign of the latter. Thus, the product of (45) and (51) becomes

$$\dot{A}\dot{B} = AB\underline{/\theta - \phi} \tag{53}$$

The proof of this operation may be had by expanding the sine and cosine of θ by Maclaurin's theorem,[1] which gives

$$\sin\theta = \theta - \frac{\theta^3}{\underline{|3}} + \frac{\theta^5}{\underline{|5}} - \frac{\theta^7}{\underline{|7}} + \cdots \tag{54}$$

$$\cos\theta = 1 - \frac{\theta^2}{\underline{|2}} + \frac{\theta^4}{\underline{|4}} - \frac{\theta^6}{\underline{|6}} + \cdots \tag{55}$$

Expanding $\epsilon^{j\theta}$ in a similar fashion gives

$$\begin{aligned} \epsilon^{j\theta} &= 1 + j\theta + \frac{(j\theta)^2}{\underline{|2}} + \frac{(j\theta)^3}{\underline{|3}} + \frac{(j\theta)^4}{\underline{|4}} + \frac{(j\theta)^5}{\underline{|5}} + \cdots \\ &= 1 + j6 - \frac{\theta^2}{\underline{|2}} - j\frac{\theta^3}{\underline{|3}} + \frac{\theta^4}{\underline{|4}} + j\frac{\theta^5}{\underline{|5}} - \cdots \\ &= 1 - \frac{\theta^2}{\underline{|2}} + \frac{\theta^4}{\underline{|4}} - \frac{\theta^6}{\underline{|6}} + \cdots \\ &\quad + j\left(\theta - \frac{\theta^3}{\underline{|3}} + \frac{\theta^5}{\underline{|5}} - \frac{\theta^7}{\underline{|7}} + \cdots\right) \end{aligned} \tag{56}$$

where $\epsilon = 2.7183$, the base of Napierian logarithms. By comparison of (56) with (54) and (55) it is evident that

$$\epsilon^{j\theta} = \cos\theta + j\sin\theta \tag{57}$$

[1] See any standard text on calculus.

Substituting (57) in (50), we get

$$\dot{A} = A\epsilon^{j\theta} \tag{58}$$

In a similar manner it may be shown that

$$\epsilon^{-j\theta} = \cos\theta - j\sin\theta \tag{59}$$

It is evident from the above that the operation of multiplication in (53) could have been written as

$$\begin{aligned}\dot{A}\dot{B} &= A\epsilon^{j\theta} \times B\epsilon^{-j\phi}\\ &= AB\epsilon^{j(\theta-\phi)}\end{aligned} \tag{60}$$

For convenience, many writers prefer the symbol $\angle$ in place of the operator ϵ^{j}. Likewise, the angles may be expressed in degrees instead of radians, as called for by (54) and (55).

Impedances are expressed in this form of notation by

$$\begin{aligned}\dot{Z} &= R + jX\\ &= \sqrt{R^2 + X^2}\,\underline{/\tan^{-1}\frac{X}{R}}\\ &= Z\underline{/\theta}\end{aligned} \tag{61}$$

If $\dot{Z} = R - jX$, the expression becomes

$$\dot{Z} = Z\overline{\backslash\theta} = Z\underline{/-\theta} \tag{62}$$

The quotient of two polar vectors is found by taking the quotient of their magnitudes and the difference of their phase angles, thus:

$$\frac{A\underline{/\alpha}}{B\underline{/\beta}} = \frac{A}{B}\underline{/\alpha - \beta} \tag{63}$$

Powers and roots are found as follows:

$$(A\underline{/\alpha})^n = A^n\underline{/n\alpha} \tag{64}$$

$$\sqrt[n]{A\underline{/\alpha}} = \sqrt[n]{A}\,\underline{/\alpha/n} \tag{65}$$

The polar form of notation greatly simplifies the processes of division, multiplication, and of obtaining roots and powers. However, polar vectors must be converted into rectangular form by means of (50) or (52) before they can be added or subtracted.

Polar notation offers the additional advantage of giving the magnitude of the current or voltage directly, whereas if these items are expressed in rectangular form the square root of the sums of the squares of the components must be obtained in order to determine the magnitude.

16. Alternating-current Power.—The instantaneous power in a circuit is given by the product of the instantaneous values of the current and voltage. It is seldom that the instantaneous power is of interest, the *average power* being usually the item of importance.

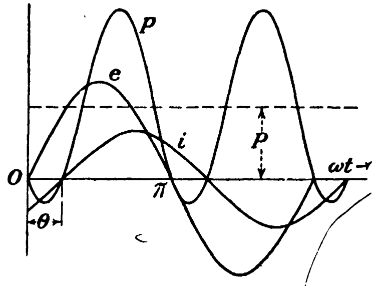


FIG. 14.—Curve of instantaneous power.

Assume a sinusoidal voltage impressed on a circuit so that the current lags behind the voltage by an angle θ. Let their equations be

$$e = E_m \sin \omega t$$
$$i = I_m \sin (\omega t - \theta)$$

The power at any instant will be

$$p = E_m I_m \sin \omega t \sin (\omega t - \theta) \tag{66}$$

Since $\sin x \sin y = \frac{1}{2} \cos (x - y) - \frac{1}{2} \cos (x + y)$, equation (66) may be written

$$p = \frac{E_m I_m}{2}[\cos \theta - \cos (2\omega t - \theta)] \tag{67}$$

The curves of e, i, and p are shown in Fig. 14. The average power will be

$$\begin{aligned} P &= \frac{1}{\pi}\int_0^\pi \frac{E_m I_m}{2}[\cos \theta - \cos (2\omega t - \theta)]d\omega t \\ &= \frac{E_m I_m}{2} \cos \theta = \frac{E_m}{\sqrt{2}} \times \frac{I_m}{\sqrt{2}} \cos \theta \\ &= EI \cos \theta \end{aligned} \tag{68}$$

where E and I are effective values. The term $\cos \theta$ is called the *power factor* of the circuit and is given by

$$\cos \theta = \frac{R}{\sqrt{R^2 + X^2}} \tag{69}$$

The product of E and I alone is often termed the *apparent power*. It becomes the true power in the absence of any reactance in the circuit.

In a circuit composed of pure reactances, $\cos \theta$ becomes zero. Consequently no power will be consumed by a circuit wherein the resistance is zero. When R is zero, the angle θ will be 90 degrees, and the curve of instantaneous power in Fig. 14 will be symmetrical about the time axis. The positive loops in this case represent power stored in the circuit. The negative loops represent the stored power being returned to the source. It will be noted that the power curve has a frequency double that of the voltage or current. This double-frequency phenomenon manifests itself in various types of electro-acoustic devices, causing them to produce outputs of twice the frequency of the applied voltage unless proper precautions are taken.

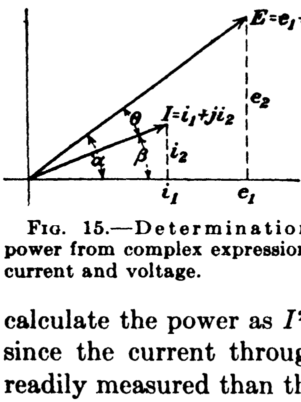


FIG. 15.—Determination of power from complex expressions of current and voltage.

In electrical communication it is usually more convenient to calculate the power as I^2R (see Sec. 4), rather than as $EI \cos \theta$, since the current through the power-absorbing device is more readily measured than the voltage across it.

When the voltage and current are expressed in complex form, the power consumed by the circuit is given by the algebraic sum of the respective products of the in-phase and quadrature components of the current and voltage. In Fig. 15 let

$$E = e_1 + je_2$$
$$I = i_1 + ji_2$$

The power will be

$$\begin{aligned} P &= EI \cos \theta = EI \cos (\alpha - \beta) \\ &= EI(\cos \alpha \cos \beta + \sin \alpha \sin \beta) \\ &= EI\left[\frac{e_1}{E} \times \frac{i_1}{I} + \frac{e_2}{E} \times \frac{i_2}{I}\right] \\ &= e_1 i_1 + e_2 i_2 \end{aligned} \tag{70}$$

If any of the components of E or I are negative, the negative sign must be used with the particular component, so that in general

$$P = (\pm e_1)(\pm i_1) + (\pm e_2)(\pm i_2) \tag{71}$$

While the power has been determined in (71) from vector expressions for the current and voltage, it is not a vector quantity. Accordingly, the total power consumed in a complicated network will be the arithmetical sum of the powers consumed in the various branches.

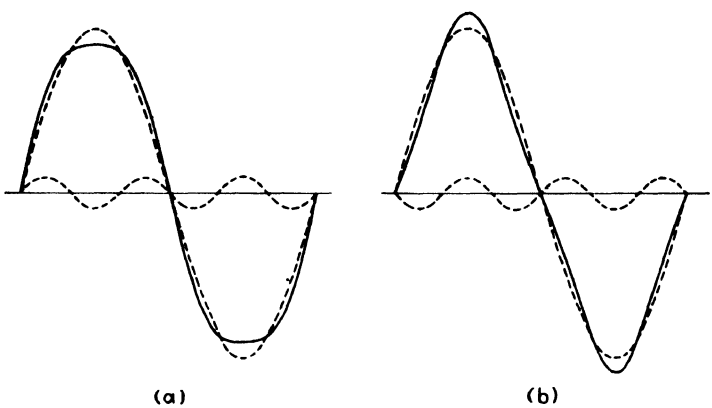


FIG. 16.—Resolution of flat-topped and peaked waves into a fundamental and third harmonic.

17. Nonsinusoidal Waves.—Many of the wave forms encountered in the field of electrical communication are far from being sinusoidal. Distorted periodic waves of this character may be regarded as being composed of a series of sine waves of various amplitudes and with frequencies which are in the ratios of 1, 2, 3, 4, 5, etc., to the frequency of the fundamental period of the observed wave. The resolution of a nonsinusoidal periodic function into a series of sine waves of multiple frequency is called a Fourier series. The first term of the series which has the same period as the observed wave is called the fundamental; the higher frequency terms are called harmonics. Theoretically, an infinite number of terms would be required to represent the average distorted wave of current or voltage, but for most

practical purposes the first few terms are usually sufficient. Such a series has the form

$$e = E_0 + E_1 \sin(\omega t + \theta_1) + E_2 \sin(2\omega t + \theta_2) + E_3 \sin(3\omega t + \theta_3) + \cdots \quad (72)$$

The initial term E_0 represents a direct-current component which may be present if partial rectification has taken place. In the case of an alternating-current generator E_0 would be zero.

Thus, the flat-topped wave of Fig. 16*a* and the peaked wave of Fig. 16*b* may be resolved into a fundamental and third harmonic of the amplitudes and phase relations shown. If the wave is symmetrical about the time axis, *i.e.*, if the wave has exactly similar positive and negative loops, it cannot contain even harmonics. The effect on symmetry of an even harmonic is shown in Fig. 17. This characteristic wave shape is frequently produced by thermionic oscillators and amplifiers. Even harmonics are absent in the wave shapes of commercial alternators. In order for these machines to produce even harmonics it would be necessary for the flux distribution under all the north poles to be different from that under the south poles, and for the armature conductors to be concentrated in a single slot. The symmetry of the field structure and the distributed nature of the armature winding prevent this, although odd harmonics of appreciable magnitude are sometimes produced as shown in Fig. 1.

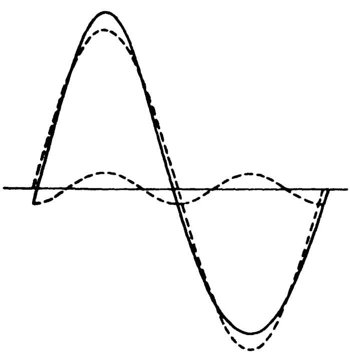


FIG. 17.—Wave shape resulting from a fundamental and second harmonic.

The Fourier series for a rectangular wave is given by

$$e = E \sin \omega t + \tfrac{1}{3}E \sin 3\omega t + \tfrac{1}{5}E \sin 5\omega t + \tfrac{1}{7}E \sin 7\omega t + \cdots \quad (73)$$

The manner in which the wave form given by this equation approaches that of a rectangular wave, as successive terms are added, is illustrated in Fig. 18. While an infinite number of terms would be needed in the series to represent a rectangular

wave exactly, it will be observed that the first four terms result in a fair approximation.

The effective or r.m.s. value of a nonsinusoidal voltage in terms of its harmonics is given by

$$E = \sqrt{\frac{E_{m_1}^2 + E_{m_2}^2 + E_{m_3}^2 + \cdots}{2}} \tag{74}$$

$$= \sqrt{E_1^2 + E_2^2 + E_3^2 + \cdots} \tag{75}$$

where the subscript m denotes maximum values. The terms without the subscript m are effective values. The effective value of a nonsinusoidal current will be of identical form, substituting I for E in either (74) or (75).

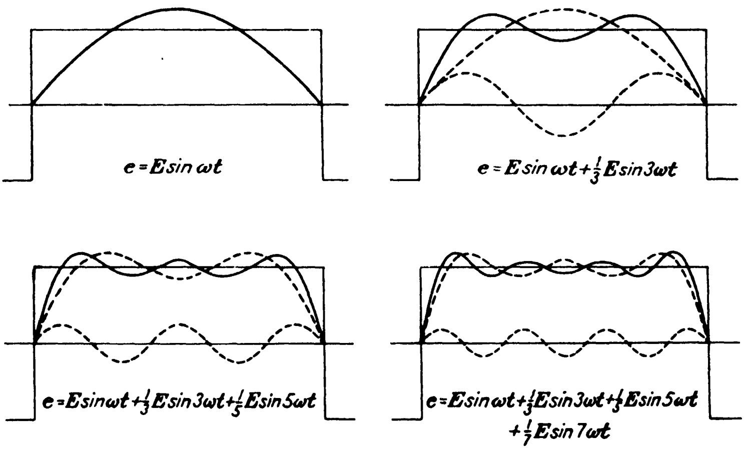


FIG. 18.—Development of a rectangular wave from Fourier series.

The solution of a circuit with a nonsinusoidal voltage impressed resolves itself into the computation of the current due to each individual voltage harmonic. If the circuit contains inductive or capacitive reactance, the resulting current wave will no longer have the same wave shape as the impressed voltage wave. The circuit impedance will be different for each harmonic, so that the amplitude ratios of the harmonic currents will now differ from those of the impressed voltage. Furthermore, the phase angles between the currents and voltages will progressively change. Since the shape of the resultant current wave is dependent upon the phase displacements of the component harmonics as well as

upon their amplitudes, a further departure from similarity results. Only when the circuit is composed of pure resistances will the current wave form be identical with that of the impressed voltage. Thus, if a complex voltage wave due to speech or music is impressed on a loud-speaker with inductive reactance predominating, the current wave through the device will no longer be identical with the impressed voltage and appreciable distortion may result.

Harmonics may be present in the current wave even if entirely absent in the applied voltage. Such a condition will occur when a voltage is impressed on a circuit whose impedance is a function of the current. An example of this is a coil wound on a closed iron core. The inductance will be μL, where μ is the permeability of the iron and L is the inductance of the coil when removed from the core. Since the permeability of the iron is a function of the current, it is evident that the instantaneous reactance of the coil is continually changing. The current wave will therefore be no longer sinusoidal.

The average power in a circuit having harmonics in both current and voltage is the sum of the powers due to each harmonic term. Let

$$e = E_{m_1} \sin (\omega t + \theta_1) + E_{m_2} \sin (2\omega t + \theta_2) + E_{m_3} \sin (3\omega t + \theta_3) + \cdots$$

and

$$i = I_{m_1} \sin (\omega t + \theta_1') + I_{m_2} \sin (2\omega t + \theta_2') + I_{m_3} \sin (3\omega t + \theta_3') + \cdots$$

The average power is

$$p = \frac{E_{m_1} I_{m_1}}{2} \cos (\theta_1 - \theta_1') + \frac{E_{m_2} I_{m_2}}{2} \cos (\theta_2 - \theta_2') + \frac{E_{m_3} I_{m_3}}{2} \cos (\theta_3 - \theta_3') + \cdots \quad (76)$$

$$= E_1 I_1 \cos (\theta_1 - \theta_1') + E_2 I_2 \cos (\theta_2 - \theta_2') + E_3 I_3 \cos (\theta_3 - \theta_3') + \cdots \quad (77)$$

The letters E and I in (76) and (77) with the subscript m represent maximum values. Without the subscript m they represent effective values.

The addition or subtraction of nonsinusoidal voltages or currents is accomplished by expressing each in terms of its Fourier

series. The fundamentals and harmonics of like frequency may then be added or subtracted vectorially to give the fundamental and harmonics of the resultant wave. Adding vectorially the effective values of two nonsinusoidal currents as read by suitable ammeters to obtain the effective value of the resultant current may result in considerable error. Equivalent sine waves as obtained from instrument readings, or computed by (75), can only be added or subtracted if their wave shapes are identical and there is no phase displacement between them. Even when the wave shapes are identical, the fact that they are displaced in phase from each other will introduce an error. Suppose two currents of equal magnitude and like wave form each contain a third harmonic and that the phase angle between their fundamentals is 60 degrees. The third harmonic will be zero in the resultant wave, but if their equivalent values had been added vectorially, the third harmonic would not have canceled.

Problems

1. An e.m.f. given by the equation $e = 1000 \sin 377t$ is impressed across a rectifier in series with a noninductive resistance of 1000 ohms. The rectifier may be assumed to be ideal, offering infinite resistance to the flow of current in one direction and zero resistance to currents in the opposite direction. What is the average value of the current in the circuit? What is the average power consumed in the resistance?

2. In the above circuit, if the internal drop in the rectifier is 15 volts for all values of current in the conducting direction, what will be the average value of current in the circuit? What is the average power consumed in the resistance?

3. An impedance coil has an inductance of 0.025 henry. When a 110-volt, 60-cycle potential is impressed across this coil, the power consumed is 500 watts. What are the resistance and power factor of the coil?

4. A magnetic type of loud-speaker has a resistance of 3000 ohms and an inductance of 1.5 henrys which may be assumed constant. Connected in series with the loud-speaker is a 2-μf condenser. The entire circuit is connected across a vacuum tube which may be regarded as a source of constant e.m.f. of variable frequency, having an internal resistance of 4000 ohms. If the value of the e.m.f. is 100 volts at $\omega = 5000$, what is the voltage across the loud-speaker? Across the condenser? What power is consumed by the loud-speaker?

5. In Problem 4 what value of impressed frequency will cause the power in the loud-speaker to be a maximum? What is the maximum power? What are the voltages across the condenser and loud-speaker under this condition?

6. A condenser C is shunted by a resistance R. What are the apparent resistance and capacitance of the combination in terms of C, R, and ω?

7. An impedance coil is connected in series with a condenser of $50\mu f$ across a 220-volt, 60-cycle circuit. If the impressed voltage is assumed to be the reference vector, the complex expression for the current in the circuit is $10 + j5$ amp. What are the resistance and the reactance of the coil?

8. Two impedances, whose complex expressions are $10 + j100$ ohms and $0 - j100$ ohms, are connected in parallel. If the total current supplied is 1 amp., what is the current through each branch?

9. A condenser and a noninductive resistance are connected in parallel across a 220-volt circuit. If the complex expression for the impressed voltage is taken as $176 + j132$, the total current is $-2 + j11$ amp. What are the complex expressions for the currents in the condenser and the resistance?

10. A coil whose impedance is $10 + j100$ ohms is connected in series with a pure capacitance of 5 ohms reactance. What value of capacitance must be connected in parallel with the coil so that the power factor of the entire circuit will be unity, if the impressed frequency is 1400 kc?

11. A coil L_1 and two condensers C_1 and C_2 are connected as shown in Fig. *A*. A second circuit is composed of L_a, C_a and C_b. Determine L_a, C_a, and C_b in terms of L_1, C_1, and C_2 in order that the resultant impedances of the two combinations may be alike at all frequencies.

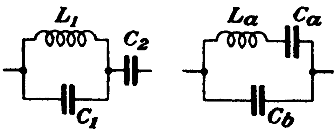


FIG. *A*.

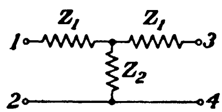


FIG. *B*.

12. An inductance L is connected in series with a capacitance C and an alternating voltage E is impressed across them. If an impedance Z is then shunted across the capacitance, show that the current through Z will be constant and independent of the nature or magnitude of Z, if $\omega^2 LC = 1$.

13. In the circuit shown in Fig. *B*, the impedance measured between terminals 1–2 is $Z_0 = 750\overline{\backslash 10^\circ}$ ohms. When terminals 3–4 are short-circuited, the impedance between 1–2 becomes $Z_s = 800\overline{\backslash 20^\circ}$ ohms. What are the values of Z_1 and Z_2?

14. The plate current of a vacuum-tube oscillator is given by

$$i = 0.552 + 0.96 \cos \omega t + 0.543 \cos 2\omega t + 0.14 \cos 3\omega t - 0.07 \cos 4\omega t - 0.105 \cos 5\omega t - 0.043 \cos 6\omega t.$$

This current flows through a "tank" circuit composed of a coil and condenser in parallel. The resistance associated with the coil is adjusted so that the power factor of the parallel circuit is unity for the fundamental frequency of $\omega = 5 \times 10^6$. At this frequency the condenser has a pure reactance of 303 ohms. The inductance of the coil is 60 μh. What is the resistance of the coil?

15. In Problem 14, what is the maximum value of the fundamental voltage across the parallel circuit? What is the power absorbed at the fundamental frequency?

16. In Problem 14, what is the maximum value of the second harmonic voltage across the parallel circuit? What is the power absorbed at this frequency?

17. In Problem 14, what is the maximum value of the third harmonic voltage across the parallel circuit? What is the power absorbed at this frequency?

18. In Problem 14, neglecting the resistance of the coil in comparison with its reactance for all frequencies higher than the fundamental, what is the Fourier series for the voltage across the parallel circuit? What is the effective value of this voltage? What are the percentage amplitudes of the harmonic currents and voltages in terms of their fundamentals?

CHAPTER II

RESONANT CIRCUITS

18. Series Resonance.—A circuit composed of resistance, inductance, and capacitance in series is said to be in resonance when the total reactance is zero. The current in a circuit of this type will be given by

$$I = \frac{E}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}} \tag{1}$$

At resonance

$$\omega L = \frac{1}{\omega C} \qquad \text{or} \qquad \omega^2 LC = 1 \tag{2}$$

and

$$f_r = \frac{1}{2\pi\sqrt{LC}} \tag{3}$$

If the frequency of the impressed voltage is constant, resonance may be produced by adjusting either L or C; or if L and C are fixed, the frequency may be varied to secure resonance. Equation (3) indicates that for any value of L and C there will be some frequency that will produce resonance. At resonance the impedance of the circuit is a minimum and is equal to the resistance. Under this condition the resonant current is given by

$$I = \frac{E}{R} \tag{4}$$

and will be a maximum and entirely independent of the magnitudes of the inductive and capacitive reactances.

The voltage drop across the inductance will be $I\omega L$ and that across the capacitance will be $I/\omega C$. These two voltages will be equal in magnitude and 180 degrees out of phase with one

another, as shown in Fig. 19. The impressed voltage is, therefore, equal to the voltage drop across the resistance. If the resistance of the circuit is small compared with the inductive and capacitive reactances, the voltages across the latter two will greatly exceed the impressed voltage, as will be seen from the following example.

Let E in Fig. 19 be 100 volts at 60 cycles and let R, L, and C be 1 ohm, 1 henry, and 7.04 μf, respectively. Then

$$\omega L = 377 \text{ ohms}$$
$$\frac{1}{\omega C} = \frac{10^6}{377 \times 7.04} = 377 \text{ ohms}$$
$$I = \frac{E}{\sqrt{R^2 + (X_L - X_C)^2}} = \frac{100}{\sqrt{1^2 + (377 - 377)^2}} = 100 \text{ amp.}$$
$$E_R = IR = 100 \text{ volts}$$
$$E_L = IX_L = 37{,}700 \text{ volts}$$
$$E_C = IX_C = 37{,}700 \text{ volts}$$

A lower value of circuit resistance would have resulted in an even greater resonant rise in voltage across the coil and condenser.

If a coil of negligible resistance is connected in series with a condenser, and a constant voltage of variable frequency is impressed, the current will vary, as shown by the dotted curve in Fig. 20. The inductive reactance will be a straight line through the origin, while the condensive reactance will be a hyperbola. The total impedance, if the resistance is negligible, will be $Z_T = X_L - X_C$. The curve representing the current will be the reciprocal of the impedance curve since the impressed voltage is constant. The current will be infinite at the resonant frequency if the resistance is zero. In practice there will always be more or less resistance present, so that the current will be a maximum at resonance and finite in value, as shown.

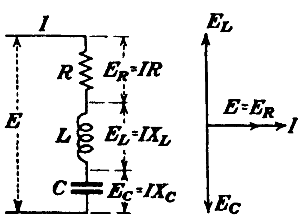


FIG. 19.—Vector diagram of a series resonant circuit.

19. Sharpness of Resonance.—The effect of resistance on the resonance curve of current is shown in Fig. 21. The greater the resistance, the broader the curve and the lower the resonant rise in the current. The majority of the resistance in the circuit

is associated with the coil, the losses in the condenser being usually negligible in comparison. The figure of merit of a coil is expressed as the ratio of the reactance to the resistance and is usually referred to by the symbol Q, so that

$$Q = \frac{\omega L}{R} \tag{5}$$

The resistance of a coil at radio frequencies increases with the frequency due to skin effect, eddy currents in the conductors, dis-

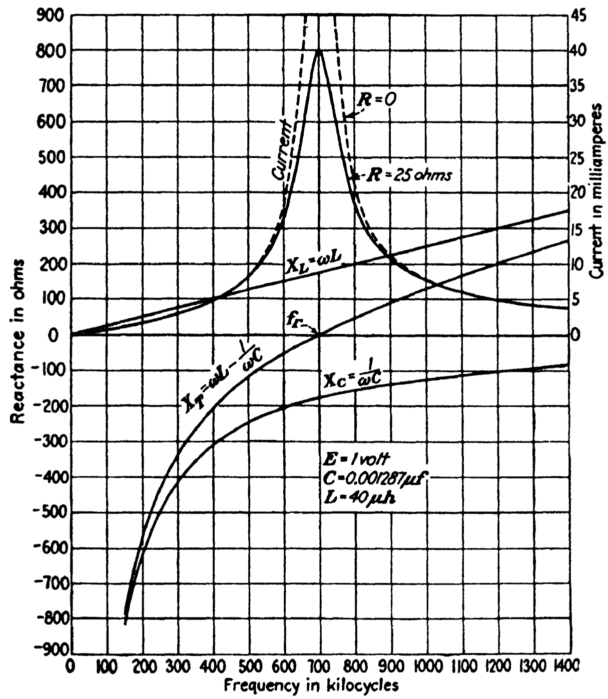


FIG. 20.—Variation of current and reactance in a series resonant circuit.

tributed capacitance between turns, and dielectric losses. These factors result in Q remaining reasonably constant for a limited range of frequencies, as shown in Fig. 22. The higher the value of Q, the sharper will be the resonance curve for the circuit. The sharpness of resonance is also dependent upon the ratio of L to C in the circuit. If the circuit resistance is held constant, the rapidity with which the current changes in value in the

vicinity of resonance will depend upon the rate of change of the total reactance with frequency. In other words, the steeper the slope of the total reactance curve X_T at the point f_r in Fig.

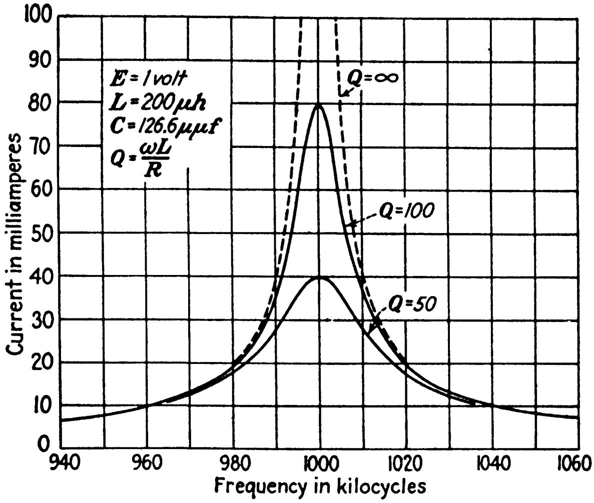


FIG. 21.—Effect of resistance on a series resonant circuit.

20, the sharper will be the resonance curve. This will be brought about by making the ratio of L to C large. Conversely, a small

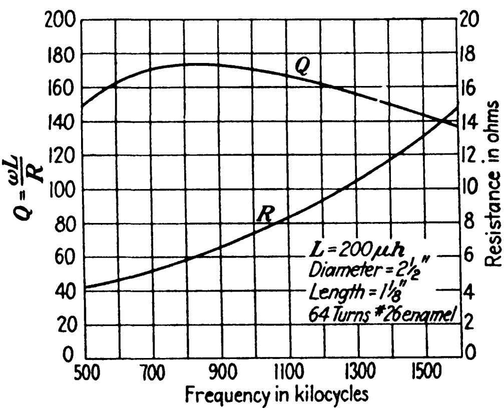


FIG. 22.—Variation of R and Q with frequency for a typical single-layer solenoid.

coil in series with a large condenser will have a relatively broad resonance curve. This assumes that the total resistance remains constant while the ratio of L to C is varied, which in reality is a

variation in Q. Increasing the inductance of a coil without increasing its resistance calls for a more efficient design of greater dimensions and weight. If Q remains constant as the inductance of the coil is varied, the sharpness of resonance is independent of the ratio of L to C. With a fixed impressed voltage and Q remaining constant, the current at resonance will progressively diminish as the LC ratio is increased, but the ratio of the current at resonance to the current at some higher or lower frequency will be unchanged.

If Q of the circuit is known, the sharpness of resonance may readily be estimated to an error of less than 1 per cent by recalling that a change of $1/Q$ in the impressed frequency will reduce the current to 45 per cent of its resonant value. Thus, with the coil of Fig. 22 which has a value of Q of 170 at 1000 kc, the current would be reduced to 45 per cent of the resonant value when the frequency is $\frac{1}{170}$ of 1000 kc, or 5.88 kc off resonance.

The resonant rise in voltage across the coil or condenser will be approximately Q times the impressed voltage. Referring to Fig. 19,

$$E_L = E_C = I\omega L = \frac{E}{R}\omega L = EQ \tag{6}$$

Equation (6) assumes that the voltage across the coil is given by $I\omega L$, whereas it is actually $I\sqrt{R^2 + \omega^2 L^2}$. This difference can usually be neglected except where Q is small. Thus, if the coil of Fig. 22 were tuned by a suitable condenser and 1 volt at 1000 kc. was impressed across the entire circuit, the voltage across the coil would rise to 170 volts at resonance.

The variation of the voltage across the coil as the frequency is varied is shown in Fig. 23, while the voltage across the condenser is shown in Fig. 24. Both of these curves are similar in shape, in the vicinity of resonance, to the current curves of Fig. 21. At frequencies considerably above resonance the voltage across the coil approaches the value of the impressed voltage, while the voltage across the condenser approaches the impressed voltage as the frequency approaches zero.

20. Parallel Resonance.—If a coil and condenser, both having negligible resistance, are connected in parallel across a source of constant voltage whose frequency is variable, the current through

the coil will be

$$I_L = \frac{E}{j\omega L} = -j\frac{E}{\omega L} \tag{7}$$

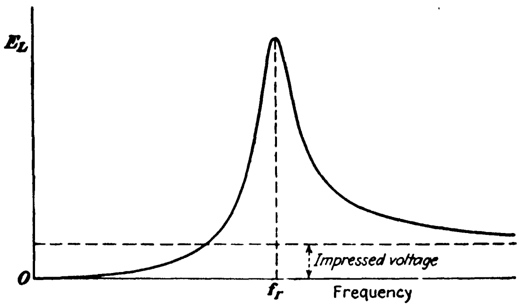


FIG. 23.—Variation of voltage across coil in series resonant circuit.

and it will lag behind the impressed voltage by 90 degrees. The current through the condenser will be

$$I_C = \frac{E}{-j\frac{1}{\omega C}} = jE\omega C \tag{8}$$

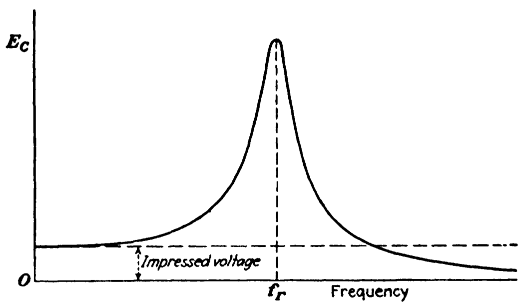


FIG. 24.—Variation of voltage across condenser in series resonant circuit.

and it will be 90 degrees ahead of the impressed voltage. The current through the coil will vary inversely with the frequency, while the condenser current will be directly proportional to the frequency, as shown in Fig. 25. The total current will be the vector sum of I_L and I_C, and will be equal to their numerical difference, since they are 180 degrees out of phase with each other. It will be observed that the variations of the currents with

frequency in this case are identical with the variations of the impedances in the series case. At resonance I_L and I_C are equal and therefore the total current is zero. The impedance of an ideal parallel resonant circuit is consequently infinite at resonance. The paradox of a current flowing in the coil and condenser while no current is being supplied by the source is explained by the fact that the circuit is now oscillating, and the condenser is alternately charging and discharging through the coil in synchrony with the impressed frequency. As the coil and condenser have no

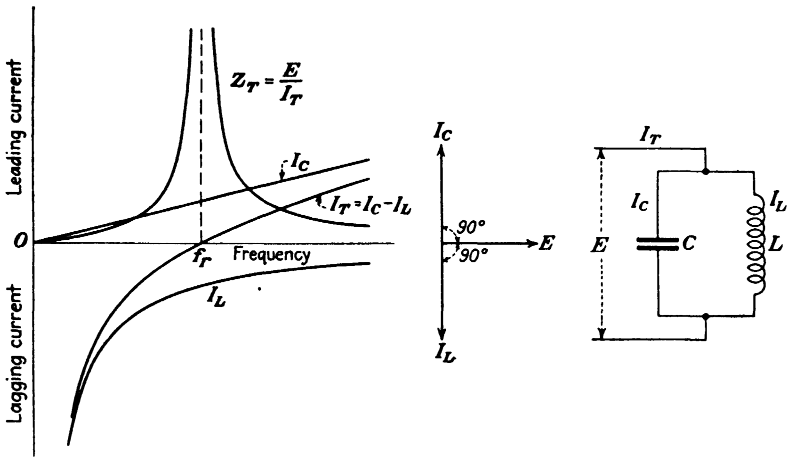


FIG. 25.—Current and vector relations in ideal parallel resonant circuit.

resistance, none of the energy represented by the initial charge taken by the condenser from the source is dissipated. The system is analogous to a frictionless pendulum which will oscillate continuously upon receiving an initial displacement without further energy being supplied.

Parallel resonance occurs in circuits of zero resistance when the currents in the two branches are equal. This results when $\omega L = 1/\omega C$, so that the expression for the resonant frequency is

$$f_r = \frac{1}{2\pi\sqrt{LC}} \tag{9}$$

which is identical with the series case given by (3). The resultant impedance of the circuit of Fig. 25 is

$$Z_r = \frac{Z_L Z_C}{Z_L + Z_C} = \frac{j\omega L\left(-j\frac{1}{\omega C}\right)}{j\left(\omega L - \frac{1}{\omega C}\right)} = -j\frac{\omega L}{\omega^2 LC - 1} \tag{10}$$

and will be of the nature of an infinite resistance when $\omega^2 LC = 1$.

21. Parallel Resonance When the Coil Contains Resistance.—In practice there will always be some resistance present in the circuit so that the currents through the coil and condenser will be less than 180 degrees out of phase. This causes the resultant impedance of the combination to be a maximum at resonance instead of being infinite. With resistance present the problem of defining parallel resonance becomes more complicated. Two criteria suggest themselves, namely, the condition for maximum impedance, and the condition for unity power factor. The latter is merely another way of stating that the circuit behaves as a pure resistance, or that the apparent reactance is zero. It will be presently shown, depending on what is considered to be the variable, that an adjustment of the circuit to produce maximum impedance does not always produce unity power factor at the same time.

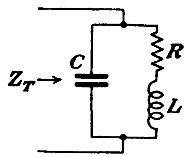


FIG. 26.—Circuit containing a perfect condenser in parallel with a coil containing resistance and inductance.

Consider the circuit shown in Fig. 26. The impedance of the combination will be the product over the sum of the two branch impedances. Hence

$$Z_T = \frac{(R + j\omega L)\left(-j\frac{1}{\omega C}\right)}{R + j\left(\omega L - \frac{1}{\omega C}\right)}$$

Rationalizing by multiplying numerator and denominator by $R - j\left(\omega L - \frac{1}{\omega C}\right)$ and collecting terms we get:

$$Z_T = \frac{\frac{R}{\omega^2 C^2} + j\left(\frac{L}{\omega C^2} - \frac{\omega L^2}{C} - \frac{R^2}{\omega C}\right)}{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}$$

$$= \frac{R + j\omega[L - C(R^2 + \omega^2 L^2)]}{\omega^2 C^2 R^2 + (\omega^2 LC - 1)^2} \tag{11}$$

The apparent resistance and reactance will be

$$R_a = \frac{R}{\omega^2C^2R^2 + (\omega^2LC - 1)^2} \tag{12}$$

$$X_a = j\omega\frac{L - C(R^2 + \omega^2L^2)}{\omega^2C^2R^2 + (\omega^2LC - 1)^2} \tag{13}$$

where

$$R_a + jX_a = Z_T$$

If X_a is zero (the condition for unity power factor) the numerator of (13) must be zero, which occurs when

$$C = \frac{L}{R^2 + \omega^2L^2} = \frac{L}{|Z_L|^2} \tag{14}$$

where Z_L is the impedance of the coil. The value of frequency which will make the apparent reactance zero, from (14) is

$$\omega = 2\pi f = \sqrt{\frac{L - CR^2}{L^2C}} = \sqrt{\frac{1}{LC} - \frac{R^2}{L^2}} \tag{15}$$

which can be written

$$\omega = \frac{1}{\sqrt{LC}}\sqrt{1 - \frac{CR^2}{L}} = \omega_0\sqrt{1 - \frac{CR^2}{L}} \tag{16}$$

where ω_0 is the resonant angular velocity of the circuit if resistance were absent. The expression for the parallel resonant frequency is no longer independent of the resistance in the circuit, as was the case with series resonance.

If $\omega^2LC = 1$, which is the condition for resonance if the resistance is zero, is substituted in (12), the apparent resistance will be

$$R_a = \frac{1}{\omega^2C^2R} = \frac{\omega^2L^2}{R} = \frac{L}{RC} \tag{17}$$

and the apparent reactance in (13) becomes

$$X_a = -j\frac{1}{\omega C} = -j\omega L = -j\sqrt{\frac{L}{C}} \tag{18}$$

For small values of R, (17) is quite large compared to (18), so that the approximate magnitude of the impedance of the parallel circuit at resonance may be obtained from (17) with sufficient accuracy for most purposes.

The exact expression for the scalar magnitude of Z_T in Fig. 26 from (11) is

$$|Z_T| = \sqrt{R_a^2 + X_a^2} = \sqrt{\frac{R^2 + \omega^2[L - C(R^2 + \omega^2L^2)]^2}{[\omega^2C^2R^2 + (\omega^2LC - 1)^2]^2}}$$

$$= \frac{\sqrt{(R^2 + \omega^2L^2)(\omega^2C^2R^2 + 1 - 2\omega^2LC + \omega^4L^2C^2)}}{\omega^2C^2R^2 + (\omega^2LC - 1)^2}$$

$$= \frac{\sqrt{R^2 + \omega^2L^2}}{\sqrt{\omega^2C^2R^2 + (\omega^2LC - 1)^2}} = \frac{|Z_L|}{\sqrt{\omega^2C^2R^2 + (\omega^2LC - 1)^2}} \tag{19}$$

If we differentiate (19) with respect to whatever is chosen as the independent variable and set the resulting equation equal to zero, the circuit relations which must exist in order that $|Z_T|$ shall be a maximum can be found. If C is the variable, differentiating (19), we obtain

$$\frac{d|Z_T|}{dC} =$$

$$\frac{-|Z_L| \times \frac{1}{2}[\omega^2C^2R^2 + (\omega^2LC - 1)^2]^{-\frac{1}{2}}(2\omega^2R^2C + 2\omega^4L^2C - 2\omega^2L)}{D^2} = 0$$

where D is the denominator of (19). Solving the above equation for C, we get

$$C = \frac{L}{R^2 + \omega^2L^2} = \frac{L}{|Z_L|^2} \tag{20}$$

which is the value of C that will make the impedance of the parallel circuit a maximum. Since (20) is identical with (14), it is evident that if C is the variable, the impedance is a maximum when the power factor of the circuit is unity. Substituting (20) in (19), the expression for the maximum impedance that can be obtained by adjusting C is

$$|Z_T|_{max} = \frac{R^2 + \omega^2L^2}{R} = \frac{|Z_L|^2}{R} \tag{21}$$

If L is taken as the variable in (19), by differentiation, $|Z_T|$ will be a maximum when

$$\omega^2L^2C - L - CR^2 = 0 \tag{22}$$

or

$$L = \frac{1 + \sqrt{1 + 4\omega^2 C^2 R^2}}{2\omega^2 C} \tag{23}$$

This relation is not the same as (20) as can be seen by solving (22) for C which gives

$$C = \frac{L}{\omega^2 L^2 - R^2} \tag{24}$$

Consequently, if L is adjusted in Fig. 26 until the circuit has maximum impedance, the power factor will no longer be unity.

Similarly, if the impressed frequency is assumed to be the variable, it is found by differentiating (19) that $|Z_T|$ will be a maximum when

$$\omega = 2\pi f = \sqrt{\frac{\sqrt{2R^2\frac{C}{L} + 1}}{LC} - \frac{R^2}{L^2}} \tag{25}$$

From the above equations it is apparent that the conditions for maximum impedance in a parallel resonant circuit containing resistance are different for all three variables C, L, and ω. The condition for unity power factor is the same for all of these variables and is given by (14) and (15). Only when C is the variable is the impedance a maximum at unity power factor. In order to avoid confusion, *parallel resonance will hereafter be taken to mean the condition of unity power factor.* The condition of maximum impedance is sometimes called *anti-resonance.* In many cases the results obtained from the two conditions are not materially different, unless Q of the coil is small. This latter condition is apt to be encountered in connection with the parallel resonant or "tank" circuits in the power amplifiers used in radio transmitters. Here the load coupled to, or shunted across, the coil frequently reduces the effective value of Q so that the adjustment for maximum impedance may be appreciably different from that for unity power factor. Furthermore, in tuning for maximum impedance a different value of maximum impedance will be obtained in each case, depending upon which element is varied in making the adjustment. This can be seen most readily by means of the vector diagram of the circuit.

Referring to Fig. 27, the vector diagram of a parallel resonant circuit is shown in the solid lines. The current I_C through the condenser leads the impressed voltage E by 90 degrees, while the current I_L through the coil lags behind E by an angle θ, where $\theta = \tan^{-1}\frac{\omega L}{R}$. The total current I_T is in phase with E so that the power factor of the circuit is unity. The impressed voltage and frequency are assumed to be constant. If C is then decreased slightly, the current through the condenser will diminish to I'_C and the resultant current will be I'_T. Increasing the value of C so that the condenser current now becomes I''_C shifts the resultant current to I''_T. As I'_T and I''_T are both larger than I_T, it follows that the impedance of the circuit is diminished if the power factor departs from unity.

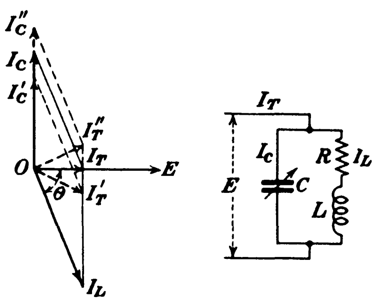


FIG. 27.—Vector diagram of parallel resonant circuit as C is varied.

If L is made the variable, the locus of I_L will be a semicircle whose diameter is E/R as shown in Fig. 28. This construction

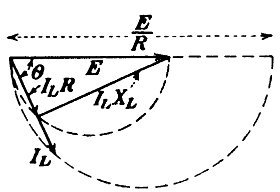


FIG. 28.—Locus of current in a circuit composed of fixed R and variable L with constant impressed voltage.

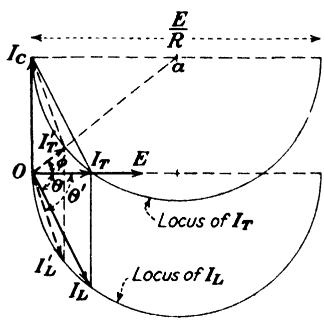


FIG. 29.—Vector diagram of a parallel resonant circuit showing locus of currents as L is varied.

is transferred to the vector diagram of Fig. 29, the heavy lines depicting the condition of unity power factor as before. The locus of I_T is obtained by shifting the circular locus of I_L vertically upward a distance of I_C, as shown. As L is varied, I_L follows

the lower circle while I_T follows the upper one. Minimum I_T, and hence maximum $|Z_T|$, occurs when I_T is perpendicular to a tangent drawn to the upper circle. At this point the total current is I'_T and leads the impressed voltage by an angle ϕ. The power factor under this condition is

$$\cos\phi = \frac{E/2R}{\sqrt{\frac{E^2}{4R^2} + I_C^2}} = \frac{E}{2R\sqrt{\frac{E^2}{4R^2} + E^2\omega^2C^2}}$$
$$= \frac{1}{\sqrt{1 + 4\omega^2C^2R^2}} = \frac{1}{\sqrt{1 + (2R/X_c)^2}} \quad (26)$$

The value of maximum impedance when L is varied can be obtained from

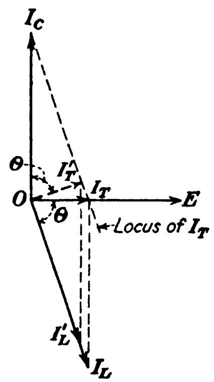


FIG. 30.—Vector diagram of a parallel resonant circuit showing locus of current if Q of the coil is constant as L is varied.

$$|Z_T|_{\max} = \frac{E}{I'_T} = \frac{E}{\sqrt{\frac{E^2}{4R^2} + I_C^2} - \frac{E}{2R}}$$
$$= \frac{2R}{\sqrt{1 + 4\omega^2C^2R^2} - 1}$$
$$= \frac{2R}{\sqrt{1 + (2R/X_C)^2} - 1} \quad (27)$$

If Q of the coil remains constant as L is varied, the angle θ will be constant and the locus of I_T will be the side of the parallelogram opposite I_L as shown in Fig. 30. As I_L is reduced to I'_L, the resultant current shifts to I'_T. The circuit impedance becomes a maximum when I_T is a minimum, which occurs when I_T is perpendicular to I_L, as at I'_T. The maximum impedance that can be obtained under this condition will be

$$|Z_T|_{\max} = \frac{E}{I'_T} = \frac{E}{I_C\cos\theta} = \frac{X_C}{\cos\theta} = X_C\sqrt{1 + Q^2} \quad (28)$$

If only R of the coil is varied, the relationships are given in Fig. 31, provided Q of the coil is greater than 1 ($\theta > 45$ degrees) when the circuit was initially adjusted to unity power factor. The locus of I_L will again be a semicircle, but the diameter is now perpendicular to E and equal to E/X_L. The locus of I_T

is a circle of the same diameter, displaced vertically upward so that the vectors I_C and I_T terminate on its circumference. As R approaches zero, I_L approaches I'_L and I_T approaches I'_T. Since the total current diminishes as R is reduced, it is apparent that the impedance of the circuit will be a maximum when R is made as small as possible. But if the initial value of Q was unity ($\theta = 45$ degrees) when the circuit was adjusted to unity power factor, $|Z_T|$ will be constant and independent of the value of R. This can be seen from Fig. 32. The resultant current I_T will be equal in magnitude to I_C regardless of the value of R so that $|Z_T|$ is constant and equal to $|X_C|$. If the initial value of Q had

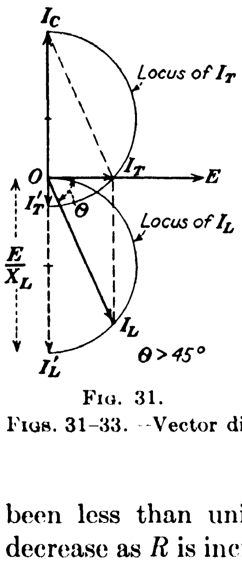


FIG. 31.

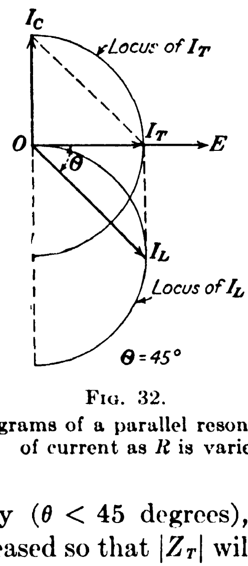


FIG. 32.

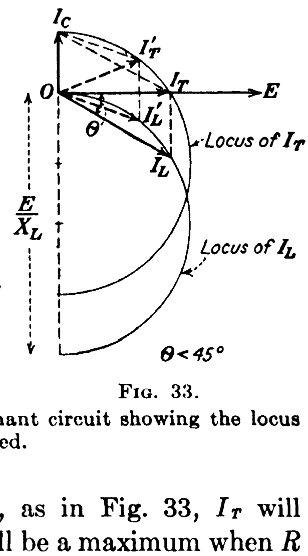


FIG. 33.

FIGS. 31–33.—Vector diagrams of a parallel resonant circuit showing the locus of current as R is varied.

been less than unity ($\theta < 45$ degrees), as in Fig. 33, I_T will decrease as R is increased so that $|Z_T|$ will be a maximum when R is made as large as possible. This is exactly opposite to the case of Fig. 31. As R approaches infinity, $|Z_T|$ approaches $|X_C|$ as a maximum value.

The effect of varying R, L, or C for maximum impedance is summarized in Fig. 34. Here it is quite evident that the value of maximum impedance obtained will be different for each variable. The power factors obtained when the impedance has been made a maximum will also differ in each case. It is easier to visualize the various relationships from the vector diagrams than from the analytical equations, so that the constructions used in Figs. 27 to 33 should be carefully studied in order to have a clear under-

standing of the behavior of parallel resonant circuits. The analytical relations for maximum impedance are summarized

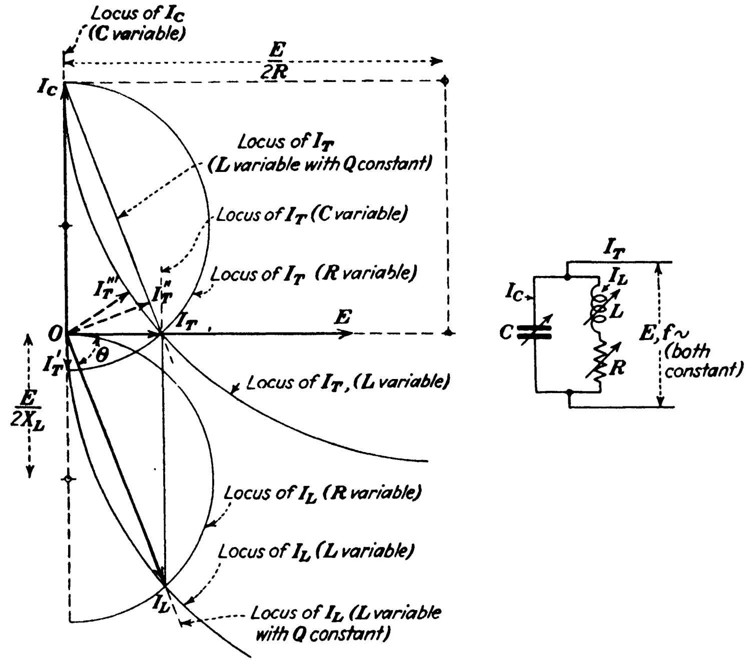


FIG. 34.—Vector diagram of a parallel resonant circuit adjusted for maximum impedance. The value of maximum impedance will depend upon which circuit element is the variable.

in Table I. Table II gives the conditions for resonance, which has been previously defined as unity power factor.

TABLE I

Variable	Condition for maximum impedance	Value of maximum impedance
C	$C = \frac{L}{R^2 + \omega^2 L^2}$	$Z_{max} = \frac{R^2 + \omega^2 L^2}{R} = R(1 + Q^2)$
L	$L = \frac{1 + \sqrt{1 + 4\omega^2 C^2 R^2}}{2\omega^2 C}$	$Z_{max} = \frac{2R}{\sqrt{1 + 4\omega^2 C^2 R^2} - 1}$
L(Q constant)	$\omega L = \frac{1}{\omega C}$	$Z_{max} = \frac{1}{\omega C}\sqrt{1 + Q^2}$
ω	$\omega = \sqrt{\frac{\sqrt{2R^2\frac{C}{L} + 1}}{LC} - \frac{R^2}{L^2}}$	$Z_{max} = \frac{L}{C\sqrt{R^2 - \frac{L}{C}\left(\sqrt{2R^2\frac{C}{L} + 1} - 1\right)^2}}$

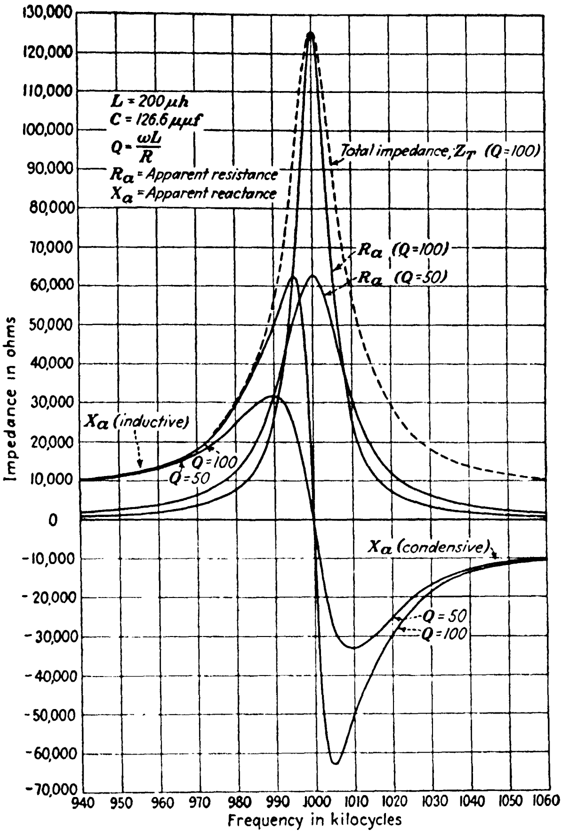


FIG. 35.—Variation of impedance with frequency in a parallel resonant circuit.

TABLE II

Variable	Condition for resonance (unity power factor)	Impedance at resonance (unity power factor)
C	$C = \frac{L}{R^2 + \omega^2 L^2}$	$Z_R = \frac{R^2 + \omega^2 L^2}{R} = R(1 + Q^2)$
L	$L = \frac{1 + \sqrt{1 - 4\omega^2 C^2 R^2}}{2\omega^2 C}$	$Z_R = \frac{1 + \sqrt{1 - 4\omega^2 C^2 R^2}}{2\omega^2 C^2 R}$
$L(Q \text{ constant})$	$L = \frac{Q^2}{\omega^2 C(1 + Q^2)}$	$Z_R = \frac{Q}{\omega C}$
ω	$\omega = \sqrt{\frac{1}{LC} - \frac{R^2}{L^2}}$	$Z_R = \frac{L}{CR}$

The variation of the apparent resistance and reactance with frequency for a typical case is shown in Fig. 35. Note that for frequencies below resonance the reactance of the circuit is inductive, which is just the opposite to the series resonant case. The sharpness of resonance is seen to be dependent upon Q. It is also dependent upon the ratio of L to C, but in a parallel resonant circuit, if R remains constant, the resonance curve becomes sharper as the ratio of L to C *is lowered.* This is again opposite to the series resonant case.

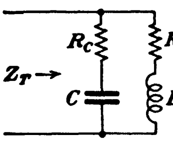


FIG. 36.—Parallel resonant circuit with the coil and condenser both containing resistance.

22. Parallel Resonance When Both Coil and Condenser Contain Resistance.—In Sec. 21 the parallel circuit considered was composed of a coil shunted by a perfect condenser. In practice, the losses in a reasonably good condenser are small so that its resistance can ordinarily be neglected. Such resistance as the condenser possesses can usually be taken into account, if necessary, by increasing the coil resistance by a sufficient amount. When the condenser resistance is appreciable, owing either to high losses or to an external load resistance in series, this method is no longer accurate.

The impedance Z_T of a circuit, such as Fig. 36, will be

$$Z_T = \frac{\left(R_C - j\frac{1}{\omega C}\right)(R_L + j\omega L)}{R_C - j\frac{1}{\omega C} + R_L + j\omega L}$$

$$= \frac{R_L + \omega^2 C^2 R_C[R_L(R_L + R_C) + \omega^2 L^2] + j\omega\{L - C[R_L^2 + \omega^2 L(L - CR_C^2)]\}}{\omega^2 C^2 (R_L + R_C)^2 + (\omega^2 LC - 1)^2} \tag{29}$$

The apparent resistance and reactance are

$$R_a = \frac{R_L + \omega^2 C^2 R_C[R_L(R_L + R_C) + \omega^2 L^2]}{\omega^2 C^2 (R_L + R_C)^2 + (\omega^2 LC - 1)^2} \tag{30}$$

$$X_a = j\omega \frac{L - C[R_L^2 + \omega^2 L(L - CR_C^2)]}{\omega^2 C^2 (R_L + R_C)^2 + (\omega^2 LC - 1)^2} \tag{31}$$

At resonance the apparent reactance will be zero, so that setting the numerator of (31) equal to zero and solving for ω,

we have for the frequency of resonance

$$\omega_r = 2\pi f_r = \sqrt{\frac{L - CR_L^2}{LC(L - CR_C^2)}}$$
$$= \omega_0\sqrt{\frac{L - CR_L^2}{L - CR_C^2}} \tag{32}$$

If the resistances of the two branches are equal, the expression for the resonant frequency becomes the same as for the series case.

The conditions for maximum impedance can be determined as before by means of the vector diagram of the circuit. The only changes in the construction will be that the current I_C through the condenser will now lead the impressed voltage by less than 90 degrees, depending on the magnitude of R_C, which will cause a corresponding shift in the loci of I_T. The various loci of I_L will be unchanged.

Problems

1. A coil having an inductance of 200 μh and a resistance of 20 ohms is connected in series with a capacitance of 0.001 μf. For what frequency of the impressed voltage will the voltages across the coil and capacitance be equal?

2. A coil having an inductance of 200 μh is connected in series with a variable condenser across a source of constant voltage and frequency. The current through the circuit is a maximum and is equal to 4 amp. when the condenser is adjusted to 126.6 $\mu\mu$f. If the capacitance of the condenser is reduced to 100 $\mu\mu$f, the current is reduced to 0.5 amp. Find the resistance of the circuit and the value of the impressed voltage.

3. A coil and condenser are connected in series across a constant voltage of variable frequency, in series with an ammeter whose deflections are proportional to the square of the current. The frequency is adjusted to resonance and the deflection of the meter is noted. The frequency is then increased to a value f_1 such that the deflection is exactly half the resonant value. If the frequency is then reduced to a value below resonance, such as f_2, so that the meter deflection is again half of the resonant value, show that the circuit resistance is given by $R = 2\pi L(f_1 - f_2)$ where L is the inductance of the coil.

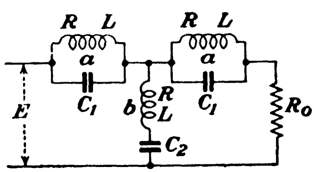


FIG. A.

4. In the filter circuit shown in Fig. A, L is 10 henrys, R is 500 ohms, and E is 100 volts at 60 cycles. The condenser C_1 is adjusted so as to make the impedance of the parallel branch a a maximum, and C_2 is adjusted so as to make the impedance of b a minimum. Find the magnitude of the voltage across R_0 if the latter is 1000 ohms pure resistance.

5. Derive an expression for the maximum impedance of a parallel resonant circuit from the vector diagram of Fig. 27, assuming C to be the variable.

6. Derive an expression for the condition of resonance (unity power factor) for the circuit of Fig. 27 in terms of ω, R, and C, under the assumption that L alone is the variable. What is the expression for the impedance at resonance in terms of these same quantities?

7. Repeat Problem 6 under the assumption that Q remains constant as L is varied. Express the condition for resonance and the resonant impedance in terms of ω, C, and Q.

8. In the circuit of Fig. 27, derive an expression for the maximum impedance in terms of ω and C if L is the variable, under the assumption that Q remains constant as L is varied. What is the expression for the maximum impedance in terms of ω, C, and Q?

9. Derive equation (25) and also the expression for the maximum impedance in terms of R, L, and C, when the impressed frequency is the variable.

10. Derive an expression for resonance (unity power factor) for the circuit of Fig. 27 in terms of R, L, and C, if the impressed frequency is the variable. What is the expression for the impedance at resonance in terms of these same quantities?

11. The tank circuit of a power amplifier is composed of a coil of variable inductance in parallel with a fixed condenser. Assuming Q of the coil to be constant and equal to 10 as L is varied, find the necessary values of L and C so that the impedance of the tank circuit will be $2000 + j0$ ohms at $\omega = 5 \times 10^6$. With C remaining fixed at the value just determined, what will be the maximum impedance that can be obtained by adjusting L? What value of L will make the impedance of the tank circuit a maximum?

12. A tank circuit composed of a coil and condenser in parallel is to have an impedance of $600 + j0$ ohms at 10^6 cycles by varying the capacitance of the condenser. The resistance of the coil is 10 ohms. Find the values of L and C needed. If C remains at this value what will be the value of the maximum impedance obtainable if L is then varied, assuming the resistance to remain constant at 10 ohms? What is this value of L?

13. A coil having a resistance of 5 ohms and an inductance of 20 μh is shunted by a condenser of 0.002 μf capacitance. What value of impressed frequency will make the impedance a maximum? What is the magnitude of the maximum impedance? What value of impressed frequency will produce resonance (unity power factor)? What is the impedance at resonance?

14. A coil of resistance R and inductance L is shunted by a capacitance C. What is the maximum value that R can possess in terms of L and C, in order for equation (25) to be valid?

15. What value of Q must a coil possess so that when tuned to parallel resonance by an ideal condenser, the impedance at resonance will be ten times the impedance of the coil alone?

16. A pure inductance L_1 and a capacitance C_1 are connected in series. Connected in parallel with them is a pure inductance L_2. If an e.m.f of variable frequency is impressed across the combination, what value of impressed frequency will make the impedance of the circuit a maximum?

$L_1 = 11.258$ μh, $C_1 = 0.001$ μf, and $L_2 = 14.072$ μh. What value of impressed frequency will make the impedance a minimum?

17. Two circuits, each composed of a coil and condenser in series, are connected in parallel with each other as shown in Fig. *B*. It is desired to make the impedance Z of the combination a maximum at 175 kc, and at the same time to have minimum impedance offered to frequencies of 165 and 185 kc. If $L_1 = 0.01$ henry, what are the necessary values for C_1, L_2, and C_2? Assume the resistance of the coils to be negligible.

18. The circuit shown in Fig. *C* is to be tuned so that the current I will be a maximum for a frequency of 10^6 cycles and a minimum for an interfering signal of 1.2×10^6 cycles. This is accomplished by tuning the coil and C_1 to parallel resonance at 1.2×10^6 cycles. Then without changing them, C_2 is adjusted so that the entire circuit between a and b is in series resonance

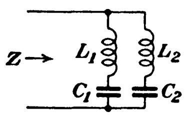


FIG. *B*

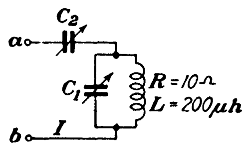


FIG. *C*.

at 10^6 cycles. Find the values of C_1 and C_2 to produce the above conditions. If the voltage impressed across a and b is 1 volt at 10^6 cycles, what will be the current I? If 100 volts at 1.2×10^6 cycles is impressed across a and b from an interfering source, what will be the current I? Compare the ratio of these two currents with the current ratio that would have been obtained if C_1 had been removed and the circuit had been tuned to series resonance at 10^6 cycles by means of C_2.

19. A coil having a resistance R_1 and an inductance L is connected in parallel with a circuit composed of R_2 and C in series. What relation must exist among R_1, R_2, L, and C in order that the total impedance of the combination shall be of the nature of a pure resistance at all frequencies? Show also that the total impedance is constant and independent of the frequency.

20. A coil whose impedance is $Z_1 = 20 + j200$ ohms is connected in parallel with an impedance $Z_2 = R_2 - jX_2$ ohms, both R_2 and X_2 being adjustable. Find the maximum value of R_2 which will still permit the circuit to be adjusted to resonance.

CHAPTER III

PROPERTIES OF COILS AND CONDENSERS

23. Inductance of a Single-layer Solenoid.—Since inductance is defined as the number of flux linkages per ampere, the calculation of the inductance of a coil resolves itself into the computation of the number of lines of magnetic force linking with the turns of the coil. This definition translated into algebra is

$$L = \frac{N\Phi}{I} \times 10^{-8} \text{ henrys} \tag{1}$$

where N is the number of turns in the coil and Φ is the magnetic flux linking with them due to the current I. The factor 10^{-8} is required to convert the expression into practical units. Only in the simplest cases can the flux linking with the turns be readily computed. The formulas for most of the commonly used types of coils are complicated functions of their dimensions so that they are usually simplified for practical computations by the use of coefficients.[1]

The flux passing through a solenoid whose length is large compared with its diameter is given by

$$\Phi = AH = A\frac{4\pi NI}{10l} \tag{2}$$

where $A = \pi r^2$, r being the radius of the solenoid in centimeters,
l = length of the winding, in centimeters.

Substituting the above in (1), the expression for the inductance becomes

$$L = \frac{4\pi^2 r^2 N^2}{10^9 l} \text{ henrys} \tag{3}$$

This can be written

$$L = \frac{4\pi^2 r^2}{10^9}\left(\frac{N}{l}\right)N \tag{4}$$

[1] *Bur. Standards Circ.* 74, also *Sci. Paper* 169.

where N/l is seen to be the number of turns per unit of length. It is apparent that the inductance is directly proportional to the number of turns in a long single-layer coil if the turns per unit of length are constant. But if length of the winding is held constant as the turns are varied, by changing the wire size or by using two or more layers, the inductance will vary as the square of the number of turns.

As the majority of single-layer coils used in radio communication have lengths not much greater than two or three times their diameters, a correction factor must be applied to equation (3) which becomes

$$L = \frac{4\pi^2 r^2 N^2}{10^9 l} K \tag{5}$$

where K is a function of $2r/l$ and is given by the curve in Fig. 37.

An approximate empirical formula due to Wheeler[2] for a single-layer solenoid is

$$L = \frac{r^2 N^2}{9r + 10l} \text{ microhenrys} \tag{6}$$

where r and l are the mean radius and length of winding in inches, respectively, and N is the total number of turns. For coils of the proportions ordinarily used in radio work, equation (6) will give an accuracy of about 1 per cent.

Maximum inductance for a given length of wire is obtained when the diameter is 2.45 times the coil length. This relationship is not very critical, however, and the ratio can vary from 1.5 to 4 without reducing L for a fixed length of wire by more than 3 per cent.

24. Multilayer Coils. Distributed Capacitance.—Single-layer coils are nearly always used for the higher values of radio frequencies when they are utilized as one of the tuned elements of the circuit. *Choke coils,* which are employed to introduce high impedance in a circuit and thus exclude radio-frequency currents from that portion of the circuit, frequently use multilayer windings. Here the losses caused by distributed capacitance are usually of no consequence. The principal objection to multilayer coils is that the distributed capacitance of the coil is greatly

[2] *Proc. I.R.E.*, vol. 16, p. 1398, October, 1928.

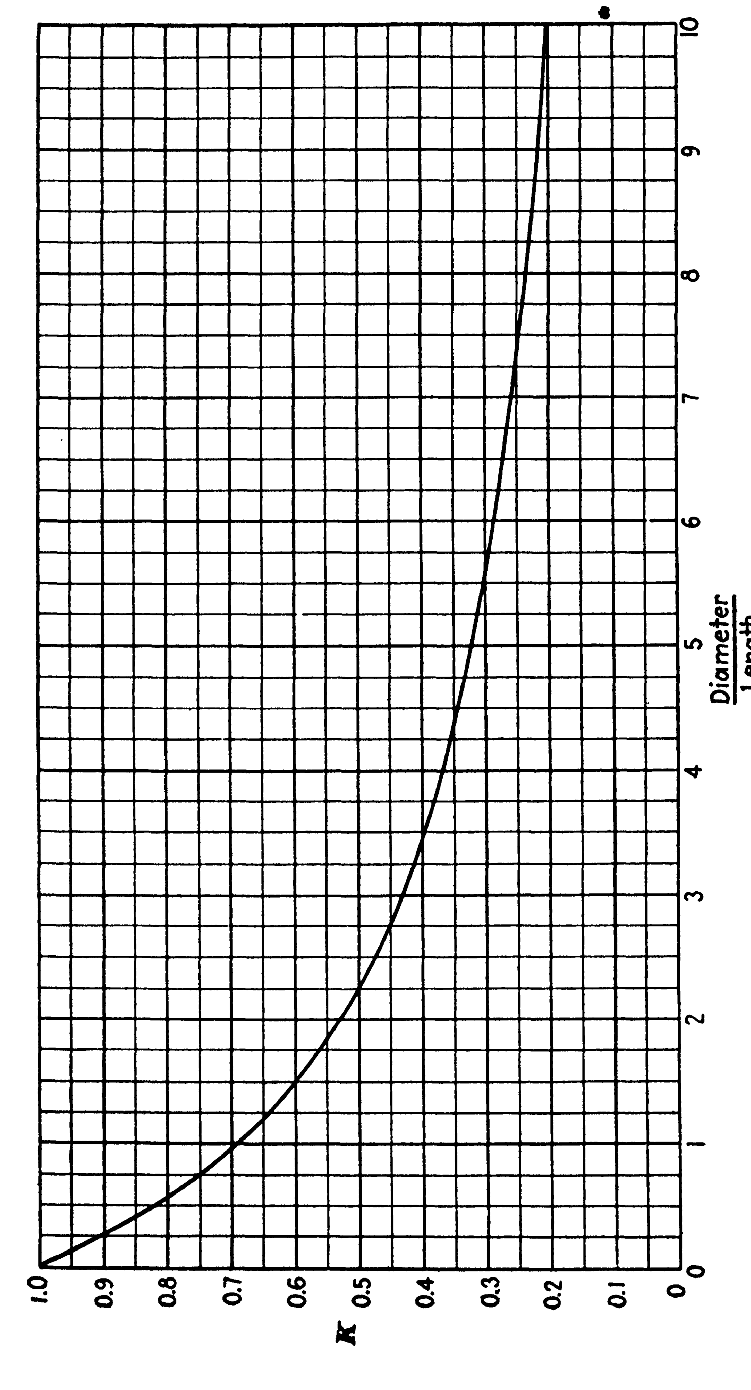


FIG. 37.—Value of K as a function of the ratio of coil diameter to length.

increased unless proper precautions are taken. In any coil there will be a small amount of capacitance between adjacent turns, between each turn and every other turn, and from each turn to ground, as shown in Fig. 38. These various individual capacitances are equivalent to a small condenser shunted across the terminals of the coil. In a multilayer coil wound in the ordinary fashion the voltage existing between the first turn of one layer and the last turn of the next layer is relatively large as compared to the voltage between adjacent turns in the same layer. This greater voltage gives rise to a larger value of displacement current through the capacitance between these turns, so that the distributed capacitance of a multilayer coil is greatly increased. If a bank winding is used, as illustrated in Fig. 39, the distributed

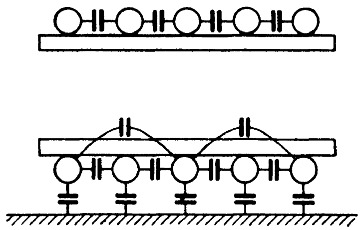


FIG. 38.

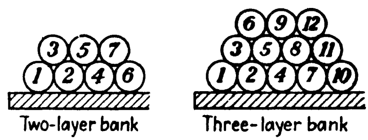


FIG. 39.

FIG. 38.—Some of the internal capacitances which contribute to the distributed capacitance of a coil.

FIG. 39.—Bank windings.

capacitance is reduced considerably below that of a similar layer-wound coil of an equal number of turns. The distributed capacitance of layer windings may be reduced by allowing a suitable space between each layer, thereby reducing the capacitance between layers. Increasing the space between the turns of each layer will also lower the distributed capacitance.

Another form of multilayer winding known as a "honeycomb" coil is illustrated in Fig. 40. These coils are layer-wound with a space between adjacent turns, as shown, and are constructed so that the wires of one layer cross those of the layer beneath at an angle, reducing the area of contact and consequently lowering the capacitance between layers. The distributed capacitance is less than would be obtained with the ordinary form of layer winding of similar dimensions. Radio-frequency choke coils commonly use a honeycomb winding.

The effect of these internal capacitances of a coil is the equivalent of a condenser shunted across the coil terminals. As a result, the coil will possess a natural resonant frequency, and at frequencies above resonance the coil will act as a condenser. At resonance the coil will have a very high value of apparent resistance, as discussed in Chap. II, and becomes particularly effec-

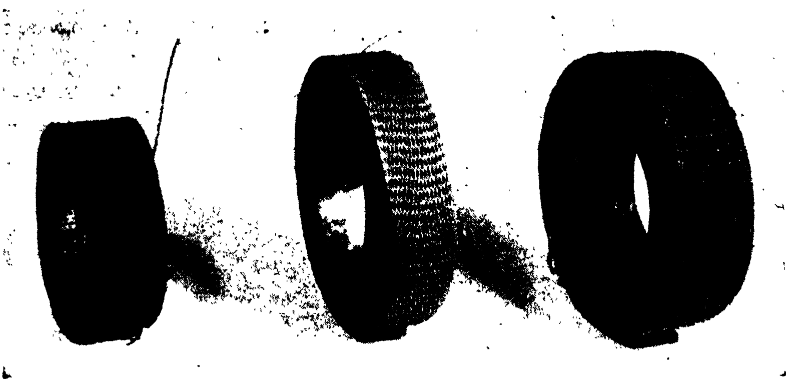


FIG. 40.—Multilayer coils of honeycomb type.

tive as a choke coil at frequencies in the vicinity of resonance. However, if used as a loading inductance L_2 in series with a smaller coil L_1, the two being tuned by a condenser as shown in Fig. 41, the high apparent resistance possessed by L_2 in the vicinity of its natural period will materially broaden the tuning and reduce the current in the circuit.

In addition to increasing the apparent resistance as the resonant frequency of the coil is approached, distributed capacitance will increase the losses in the coil. This is due to the losses in the solid dielectric in the electrostatic field between turns. The magnitude of this loss will depend upon the nature of the insulation on the wire and also upon the material of the tube on which the coil is wound. The absorption of moisture by these materials greatly increases the dielectric loss so that the coil should be suitably impregnated so as to exclude moisture. Impregnation will increase the distributed capacitance somewhat, but this will usually be more than offset by the

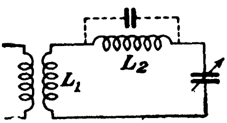


FIG. 41.—Two coils in series tuned by a variable condenser.

reduction in the losses. An impregnating material which has a low dielectric loss should be used.

A coating of collodion is satisfactory for this purpose as it also serves as a binder to hold the turns in place. Boiling the coil in paraffin, or a mixture of paraffin and rosin, is also satisfactory as this treatment tends to drive out any moisture present in the cotton or silk insulation. When collodion is used, the coil should be baked prior to its application, otherwise the moisture already present will be sealed in. Shellac and most insulating varnishes are unsatisfactory due to their larger dielectric losses and higher dielectric constants. The amount of solid dielectric in the electrostatic field of the coil should be kept as small as possible, consistent with good mechanical design, particularly at the higher frequencies. Coils for these frequencies are usually space-wound on a skeleton form or else employ a size of wire large enough to make the coil self-supporting without the use of a form. The comparatively few turns needed at the higher frequencies enable this form of construction to be practical.

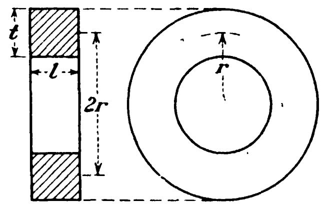


FIG. 42.—Multilayer coil of rectangular cross section.

The inductance of a multilayer coil of rectangular cross section may be computed with fair accuracy from the following empirical formula due to Hazeltine:

$$L = \frac{0.8r^2N^2}{6r + 9l + 10t} \text{ microhenrys} \tag{7}$$

where all dimensions are in inches as shown in Fig. 42; N being the total number of turns. The accuracy is greatest when r, l, and t are approximately equal.

25. Coil Resistance. Skin Effect.—In addition to the increase in coil resistance at radio frequencies caused by dielectric loss, the nonuniform distribution of the current throughout the cross section of conductor is an even greater factor. This is known as skin effect and can be best understood by considering an isolated round wire as in Fig. 43. Assuming the current to be uniformly distributed throughout the cross section of the conductor, the flux density B at a distance x from the center of the wire will be

given by the ordinate of the curve. Consider an elementary filament a at the center of the conductor. This filament will be linked by all of the flux extending from point a outward, which is proportional to the total area under the curve. A similar elementary filament b located at the surface will be linked only by the flux outside of the conductor. This flux is proportional to the area under the curve to the right of c and to the left of d. As both filaments are carrying the same amount of current the inductance of filament a will be greater than that of b since inductance is defined as the number of flux linkages per ampere.

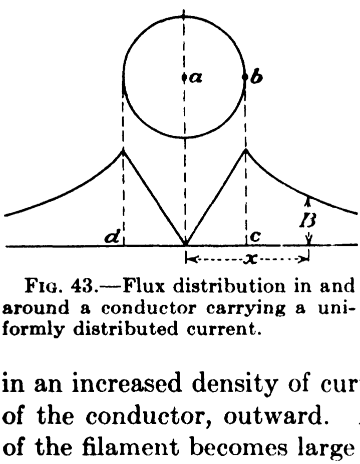


Fig. 43.—Flux distribution in and around a conductor carrying a uniformly distributed current.

If the conductor is being traversed by an alternating current, the reactance, and hence the impedance, of the central filament will therefore be greater than one located nearer to the surface. Since the wire may be thought of as being composed of a bundle of such filaments, all connected in parallel, the total current will divide inversely as the impedances of these paths. This results in an increased density of current as we proceed from the center of the conductor, outward. At radio frequencies the reactance of the filament becomes large compared to its resistance, so that practically all of the current flowing through the wire is confined to a thin shell at its surface. This distribution of the current results in a much greater I^2R loss in the conductor for a high-frequency current than for a direct current of the same value. Stated another way, the effective resistance of the conductor increases with increased frequency. Since the central portion of the wire contributes practically nothing to the conductivity at radio frequencies, copper tubing is often used in place of solid wire for coils and connections which carry considerable current. Rectangular strip is also used, although it is not so effective as generally supposed. The current is redistributed over the conductor cross section in such a manner as to be encircled by the smallest number of lines of force, so that the current density in the case of rectangular strip is greatest at the two outer edges. The remainder of the strip carries a comparatively small amount

of current, so that the effective resistance is relatively high in spite of the large amount of conductor surface. It is evident from this example that it is not the amount of surface area that determines the high-frequency resistance of a conductor, but rather the way in which the conductor material is arranged.

The ratio of the alternating-current to the direct-current resistance becomes less as the diameter of the wire is reduced. The effective resistance of a No. 39 A.W.G. ($d = 0.003531$ in.) copper wire at 1000 kc is 1 per cent higher than its direct-current resistance. A No. 18 wire ($d = 0.0403$ in.) would have an increase in resistance of 420 per cent at this frequency. These figures apply to a straight wire. If the wire had been wound in the form of a coil, the skin effect would have been much greater. Skin effect is much less in materials or alloys of high specific resistance, such as Advance or Nichrome. This is due to the impedance of the central filament being now composed chiefly of resistance, reactance no longer predominating, so that the current is more uniformly distributed throughout the cross section of the conductor. Thus, in the case of Advance, which has a resistivity about thirty times that of copper, the size of the wire could be increased to No. 23 ($d = 0.02257$ in.) before the resistance at 1000 kc became 1 per cent greater than the direct-current resistance.

Skin effect can be greatly reduced by the use of a special type of stranded cable called *Litzendraht* (usually shortened to "litz"). This is composed of a number of strands of No. 38 enameled wire transposed so that each strand is on the outer surface of the cable as much as every other strand. If this is properly done, each strand will be linked with as many lines of force as every other strand, resulting in a fairly uniform current distribution throughout the conductor. To be entirely satisfactory the strands should be woven or braided so as to be completely transposed; merely twisting them together causes a relatively small improvement. The superiority of coils wound with commercial litz over those using solid wire is confined to frequencies below 1500 kc.

26. Properties of Iron-core Coils.—The use of iron-core coils in radio apparatus is chiefly confined to audio-frequency devices such as interstage transformers and choke coils. Their use at radio frequencies is usually confined to values below 150 kc, unless especially prepared iron is employed for the core. W. J.

Polydoroff[3] describes the construction of coils for broadcast reception wherein tuning is accomplished by moving a core of molded iron dust inside of the coil, the latter being shunted by a fixed capacitance. A value of $Q = 140$ is claimed.

The chief difficulty with iron cores at the higher frequencies is due to the increase in core loss (principally eddy current) which greatly increases the apparent resistance of the coil. What might be termed the *effective permeability* of the iron also diminishes with increased frequency. This is brought about by the failure of the magnetic flux to penetrate very deeply into the laminations or particles of iron composing the core, owing to the shielding effect of the eddy currents. If we consider a particle of iron traversed by an alternating magnetic flux, eddy currents will be induced which will flow around the outer periphery of the particle in a direction such as to weaken the flux that produced them. Since the demagnetizing effect of the eddy currents will be greatest along the center line of the particle, the flux density there will be a minimum. The flux density will be a maximum at the surface, so that the flux distribution throughout the particle is quite similar to the distribution of current in a wire due to skin effect. As the permeability of the iron is defined as the ratio of the flux density to the magnetizing force, it is evident that this reduction in flux is equivalent to a reduction in the permeability.

The inductance of a coil with an iron core is

$$L = \mu L_0 \tag{8}$$

where μ is the permeability of the core and L_0 is the inductance of the coil prior to the insertion of the core. Owing to the nominal value of effective permeability obtainable at the higher frequencies and the large increase in the apparent resistance of the coil due to eddy-current loss with existing core materials, unless special precautions are taken, it is usually possible to obtain larger values of Q for a coil by the use of an air core.

The eddy-current loss varies as the square of the frequency and as the square of the thickness of the laminations if the flux density is uniform. It is difficult to roll sheet steel much thinner than 0.001 in., the usual thickness of laminations for radio-frequency coils, so that a core composed of iron dust is often

[3] *Proc. I.R.E.*, vol. 21, p. 690, May, 1933.

used. The material is mixed with a suitable binder which serves to insulate the particles from each other and is then molded under extremely high pressure into the desired shape. Reducing the eddy-current loss in this manner automatically increases the flux penetration and consequently increases the effective permeability. The flux density at which the iron is usually operated is apt to be low owing to the small value of radio-frequency voltage ordinarily impressed across the coil. The relations which exist between the voltage applied to the coil and the magnetic flux in the core may be obtained as follows: Assume the flux in the core to vary as a sine function of time so that at any instant

$$\varphi = \Phi_m \sin \omega t \tag{9}$$

The voltage induced in the coil will be

$$e = -N\frac{d\varphi}{dt} \times 10^{-8} \text{ volts} \tag{10}$$

where N is the number of turns in the coil.

Substituting (9) in (10),

$$e = -\omega N\Phi_m \cos \omega t \times 10^{-8} \tag{11}$$

The voltage will be a maximum when $\cos \omega t = 1$, so that

$$E_m = -\omega N\Phi_m \times 10^{-8} \tag{12}$$

The impressed voltage is opposite to the induced voltage so that reversing the sign of (12) and dividing by $\sqrt{2}$ to obtain the effective value, we get for the effective value of the impressed voltage,

$$E = 4.44 fN\Phi_m \times 10^{-8} \tag{13}$$

which is an equation of fundamental importance in transformer design. Consequently, with a coil or transformer of a given number of turns and area of core, the flux will vary directly with the impressed voltage, assuming constant frequency.

At low values of flux density the permeability of iron is low, having a value of only a few hundred. The maximum permeability will ordinarily be from 6000 to 10,000, depending on the chemical composition and heat-treatment of the material. Magnetization and permeability curves for a high grade of trans-

former iron are given in Fig. 44. The magnetizing force H in gilberts per centimeter is given by

$$H = \frac{4\pi NI}{10l} \tag{14}$$

where NI/l is the number of ampere-turns per centimeter. The ampere-turns per inch will be $2.02H$.

By definition,

$$\mu = \frac{B}{H} \tag{15}$$

so that the point of maximum permeability may be readily determined from the B-H curve by drawing a line from the origin

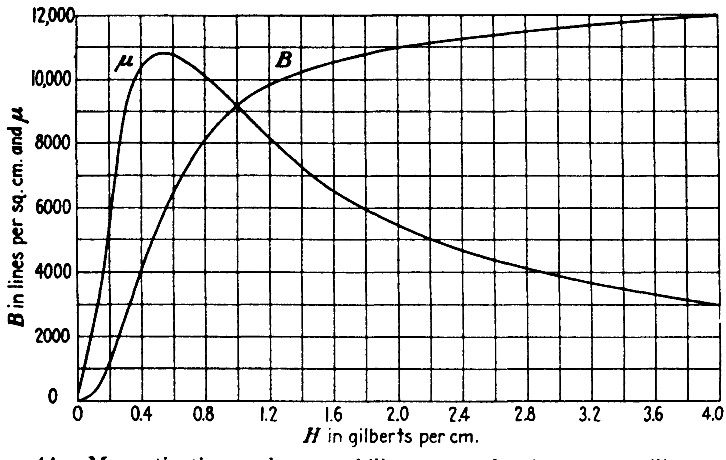


FIG. 44.—Magnetization and permeability curves for 4 per cent silicon steel

tangent to the curve. The variation of the inductance with current, from (8), will therefore be a curve identical in shape with the permeability curve in Fig. 44. This will be true if the magnetizing current is either direct or alternating, although in the latter case the inductance will vary throughout the cycle. This causes the instantaneous impedance to vary, resulting in a distorted current wave, even though the voltage impressed across the coil is sinusoidal. However, the flux in the core will continue to vary sinusoidally so that the voltage induced in the secondary coil (assuming the device to be a transformer) would be free from

distortion, notwithstanding the distorted primary current. If there is an appreciable resistance in series with the primary of the transformer, such as the internal resistance of a vacuum tube, and a sine wave of voltage is impressed across the two in series, the *IR* drop across the resistance will be nonsinusoidal because of the distorted current wave. Subtracting this nonsinusoidal *IR* drop from the impressed voltage results in a distorted voltage wave across the primary, and consequently a similarly distorted wave will be induced in the secondary.

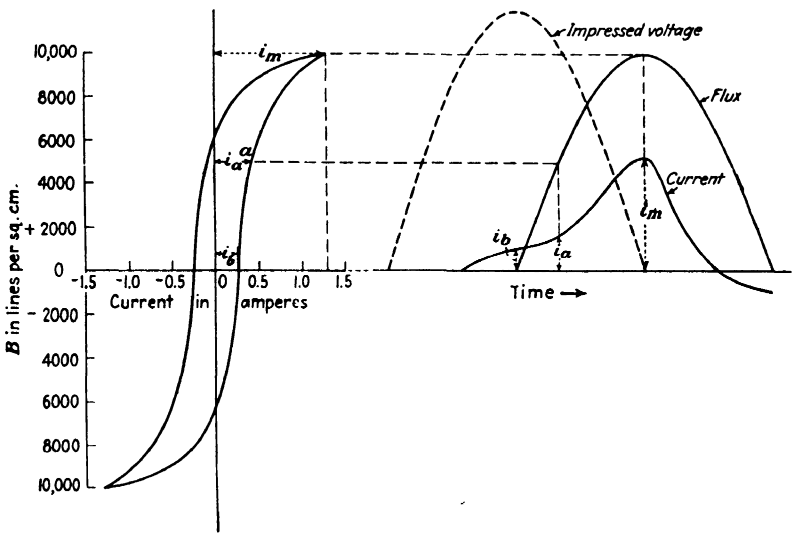


FIG. 45.—Distortion of the current wave in a coil with an iron core due to hysteresis.

The shape of the current wave can be readily obtained from the hysteresis loop of the iron, as shown in Fig. 45. This hysteresis loop is for the 4 per cent silicon steel of Fig. 44. Assuming the flux to be a sine function of time, the impressed voltage will, from (10), be 90 degrees ahead of Φ in phase. The relationship between the flux and the magnetizing current will be given by the hysteresis loop for the iron. At any point *a* on the hysteresis loop the magnetizing current will be i_a. If point *a* is projected horizontally across to the flux wave, the point of intersection will be the instant in time when the magnetizing current has the value i_a. Other points can be obtained in a similar manner.

27. Magnetic Alloys.—The flux densities employed in the cores of choke coils and transformers used in radio circuits are often quite small. This is due to the very small currents which are frequently dealt with and also to the necessity of designing the apparatus so as to keep down the core loss. These conditions require a material having a high initial permeability, low hysteresis loss, and a high specific resistance so as to keep the eddy-current loss small. The initial permeability is defined as the value of the slope of the *B-H* curve at the origin.

An alloy developed by the Bell Telephone Laboratories of 78.5 per cent nickel and 21.5 per cent iron known as *permalloy*[4] was the first to fulfill these requirements. It possesses rather remarkable magnetic properties at low values of magnetizing force. By proper heat-treatment it can be given an initial permeability of almost 10,000 and a maximum permeability of over 100,000 at a flux density of about 5000 lines per square centimeter. Its hysteresis loss is much less than for the best commercial silicon steel. There is, however, one important limitation to the material, namely, it is very susceptible magnetically to mechanical strains. Stresses below the elastic limit may reduce the initial or maximum permeability to a very small percentage of the unstressed values. It is usually necessary to shape the material to its final form and then heat-treat it in that condition.

The specific resistance of 78.5 permalloy is about 16 microhms per centimeter cube as against 56 microhms for 4 per cent silicon steel and 10 microhms for pure commercial iron, so that the eddy-current loss will be greater than for silicon steel of the same thickness of laminations. The magnetic properties are obtained by a heat-treatment which is different from that which gives the most favorable results for other magnetic materials. This is probably the reason why this alloy remained so long undiscovered. In addition to an ordinary anneal, 78.5 permalloy requires a special heating at 600°C., followed by a rapid cooling.

It was found that if some of the iron were replaced by 3.8 per cent of either molybdenum or chromium, the nickel content remaining at 78.5 per cent, a simple anneal sufficed to develop the desired magnetic properties. Magnetization curves for these three permalloys are shown in Fig. 46 and the corresponding

[4] H. D. ARNOLD and G. W. ELMEN, Permalloy, an Alloy of Remarkable Magnetic Properties, *Jour. Frank. Inst.*, vol. 195, p. 621, May, 1923.

permeability curves in Fig. 47. As will be observed, the addition of these nonmagnetic materials has lowered the maximum permeability, but the initial permeability has been increased, being 21,500 in the case of the molybdenum alloy. The specific resistance is also greatly increased, resulting in a reduction of the eddy-current loss. The addition of chromium produces the greatest increase, giving a value of 64 microhms, while 3.8 per cent molybdenum increases the resistance per centimeter cube

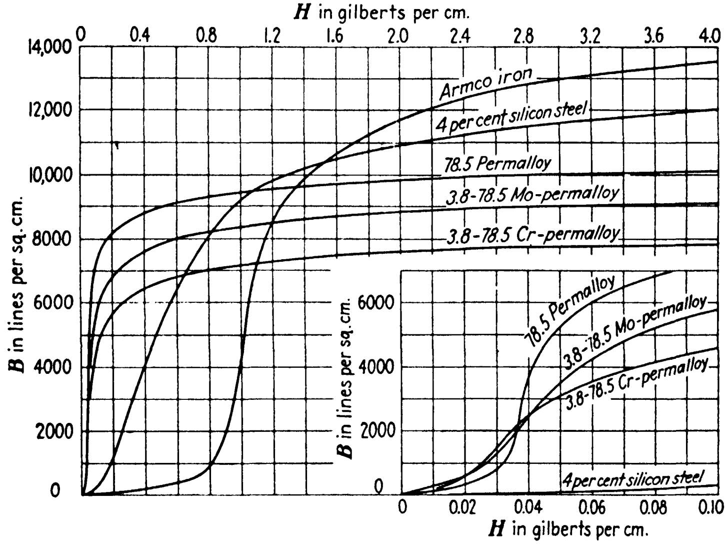


FIG. 46.—Magnetization curves of various permalloys compared with silicon steel and Armco iron.

to 57 microhms, or about the same as for 4 per cent silicon steel. These alloys are frequently used in the form of dust cores,[5] particularly in the case of loading coils for telephone lines.

The fact that permalloy approaches saturation at relatively low values of magnetizing force, in addition to its relatively high cost, restricts its applications. When used as a core material for audio-frequency transformers the direct-current component usually present in the primary winding is sufficient ordinarily to saturate the core to such an extent that the magnetic properties

[5] W. J. SHACKELTON and I. G. BARBER, Compressed Powdered Permalloy, *Trans. A.I.E.E.*, vol. 47, p. 429, February, 1928.

of the alloy may actually be poorer than in ordinary grades of iron under similar conditions.

An alloy of 50 per cent nickel-iron is often used for high-grade audio-frequency transformers. When properly annealed, initial permeabilities of 3000 can be obtained, with corresponding maximum permeabilities of 50,000 or more. It has the further advantage of possessing a higher saturation value than permalloy, although still below that of ordinary silicon steel.

If an alloy of 50 per cent nickel-iron is subjected to a prolonged heating in an atmosphere of hydrogen at a temperature of 1000° to 1200°C. a remarkable improvement results in its magnetic

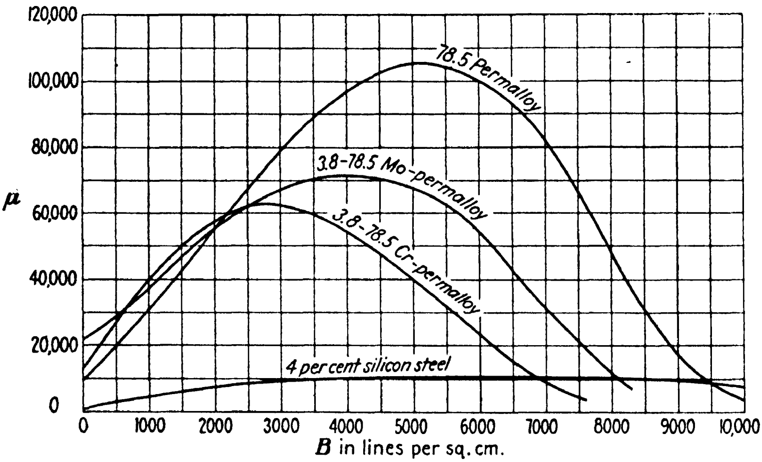


FIG. 47.—Permeability curves of various permalloys compared with silicon steel.

properties. This alloy is known as *hypernik*. It has an initial permeability which approaches that of permalloy and maximum permeability, depending on the heat-treatment, of from 60,000 to 160,000 at flux densities of about 5000 lines per square centimeter. At the higher values of B the permeabilities of the other alloys fall below that of silicon steel. With hypernik the permeability continues to remain higher than silicon steel up to flux densities of 16,000 lines per square centimeter so that difficulties due to saturation at low magnetizing forces are absent. A similar heat-treatment of ordinary iron in hydrogen likewise results in a large increase in the permeability. Initial and maximum values of permeability of 4000 and 180,000, respec-

tively, have been obtained, as compared with 250 and 9000 for the annealed iron prior to hydrogenization.

Another group of alloys composed of iron, nickel, and cobalt, known as *perminvar*,[6] are of importance in that the permeability, with proper heat-treatment, remains constant for an appreciable range of magnetizing forces. Permeability curves for various heat-treatments of an alloy having 45 per cent nickel, 25 per cent

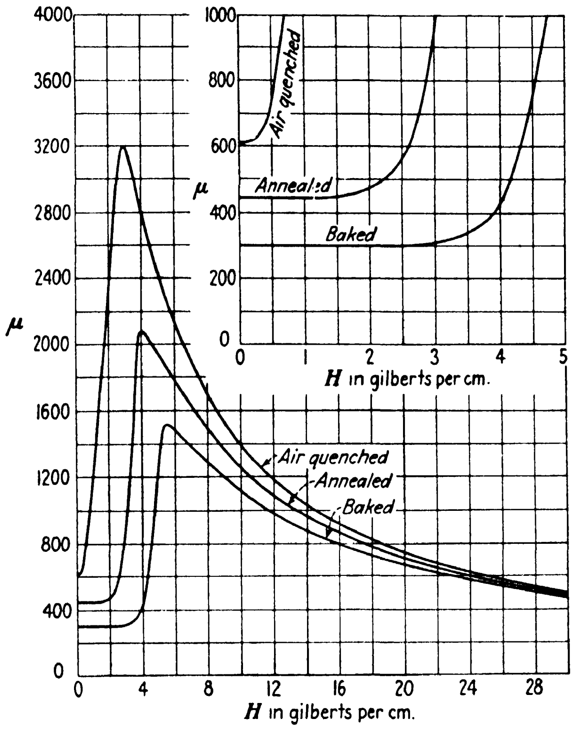


FIG. 48.—Permeability curves for perminvar.

cobalt, and 30 per cent iron are shown in Fig. 48. The sample marked "baked" was first annealed and then held at 425°C. for a considerable time. The initial permeability is somewhat higher than for ordinary silicon steel and in the case of the baked sample it remains constant for values of H up to 2.5. For this same range of magnetizing force the permeability of the silicon steel of Fig. 44 rises to a maximum of 10,800 and then falls off to a value of 4400.

[6] G. W. ELMEN, Magnetic Alloys of Iron, Nickel, and Cobalt, *Jour. Frank. Inst.*, vol. 207, p. 583, May, 1929.

In addition to the constant-permeability feature, the hysteresis loss for flux densities within the range where μ remains constant is extremely small. This is illustrated in Fig. 49 where the upper halves of hysteresis loops of various magnetic materials are shown for a maximum flux density of 600 lines per square centimeter. The loop for perminvar has no measurable area, but other more sensitive methods show its hysteresis loss to be of the order of 0.1 or 0.2 per cent of that of permalloy. For flux densities above those for which μ remains constant the hysteresis loop of perminvar begins to acquire an appreciable area as shown in Fig. 50. At a maximum flux density of 5000 the hysteresis loss of the baked sample is very much larger than for permalloy.

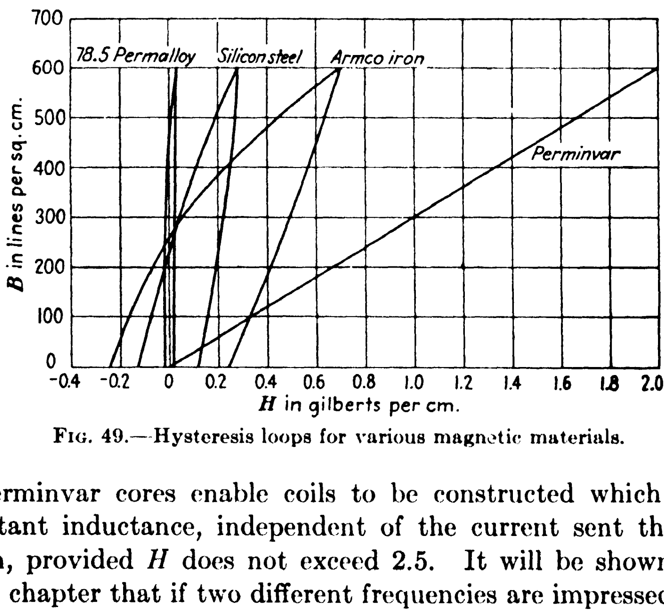


FIG. 49.—Hysteresis loops for various magnetic materials.

Perminvar cores enable coils to be constructed which have constant inductance, independent of the current sent through them, provided H does not exceed 2.5. It will be shown in a later chapter that if two different frequencies are impressed on a nonlinear impedance—one which does not follow Ohm's law—modulation will occur. This means that new frequencies will be produced which were not present in the applied voltages, and that the imprint of one frequency is left on the other. This problem is present in carrier-current telephony where several different speech-modulated high-frequency voltages are impressed on the primary of the same transformer. The use of a perminvar core with its linear hysteresis loop will prevent cross talk and interference due to these modulation effects.

28. Incremental Permeability.—In many cases the iron-core transformers and choke coils used in radio circuits carry a continuous current with a smaller alternating-current ripple superimposed upon it. In these cases we are almost always concerned with the amount of inductance offered to the alternating-current component.

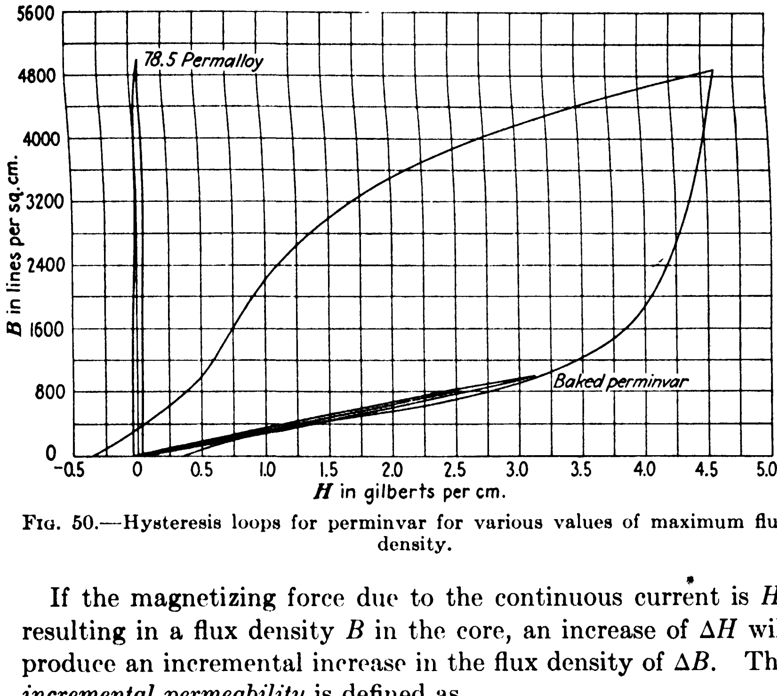


FIG. 50.—Hysteresis loops for perminvar for various values of maximum flux density.

If the magnetizing force due to the continuous current is H, resulting in a flux density B in the core, an increase of ΔH will produce an incremental increase in the flux density of ΔB. The *incremental permeability* is defined as

$$\mu_\Delta = \frac{\Delta B}{\Delta H} \tag{16}$$

which would be proportional to the slope of the magnetization curve at the point in question if it were not for the effect of hysteresis. Thus, if the magnetizing force is reduced by an amount ΔH, the corresponding reduction in the flux will not follow the magnetization curve, but will lie above it, following the path of a hysteresis loop.

When a small alternating-current ripple is superimposed on a continuous magnetizing force, as in Fig. 51, a displaced hysteresis

loop will be produced and the values of ΔB and ΔH called for in (16) must be measured from the tips of the displaced hysteresis loop. The continuous magnetizing force is H_0, which produces a flux density of B_0 in the core. The quantity ΔH is proportional to the double amplitude of the superimposed alternating current. Since the inductance is directly proportional to the permeability, the inductance offered to the continuous current is proportional to the slope of the line Oc, while that offered to the alternating-current component is proportional to the slope of ab, which is much less. Furthermore, the incremental permeability steadily diminishes as the direct-current flux density is increased,

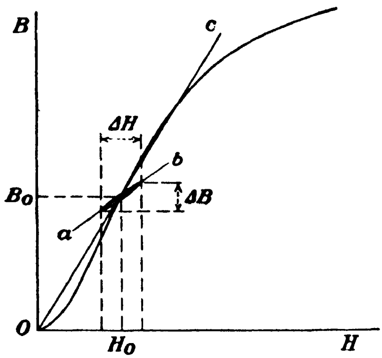


FIG. 51.—Displaced hysteresis loop caused by an alternating-current ripple superimposed on a continuous magnetizing force.

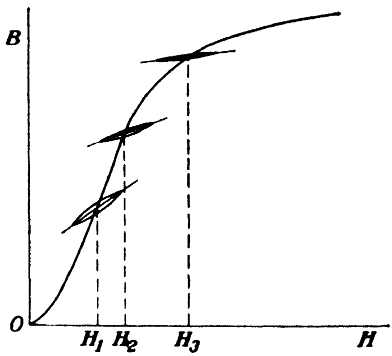


FIG. 52.—Effect of a continuous magnetizing force on incremental permeability.

as shown in Fig. 52. Increasing the amplitude of the alternating-current magnetizing force ΔH tends to increase the slope of the displaced hysteresis loop, causing an increase in μ_Δ. The incremental permeability is usually about 20 per cent of the normal permeability at low flux densities and falls off to about 10 per cent of the normal value for values of B in the vicinity of 10,000 lines per square centimeter.

This decrease in inductance due to the reduction in incremental permeability as the direct-current magnetizing force is increased is illustrated in Fig. 53. The curves show the variation in the primary inductance of a high-quality audio-frequency transformer as the continuous current in the primary is varied. The two curves are for identical transformers except for the core

material. Note that the alloy is affected to a greater extent than the silicon steel by the direct-current saturation of the core.

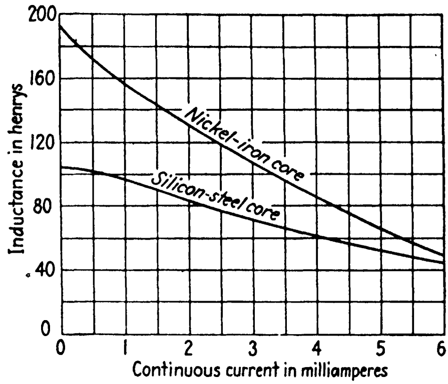


FIG. 53.—Decrease in primary inductance of an audio-frequency transformer due to continuous current in primary.

When considerable direct current must be carried by the coil, as in the choke coils of filter circuits of rectifiers, the inductance offered to the superimposed alternating current can be increased by the introduction of an air gap into the magnetic circuit. This is illustrated in Fig. 54. The best length of air gap will depend upon the core dimensions and the magnitude of the direct-current ampere-turns.[7] Air gaps are sometimes used in the magnetic circuits of audio-frequency transformers, particularly if an alloy core material is to be used.

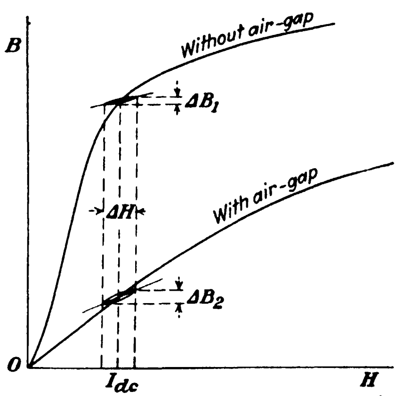


FIG. 54.—Showing the increase in incremental permeability for an iron-core coil carrying a continuous current when an air gap is introduced.

29. Inductance of Iron-core Coils with Direct and Alternating Current Superimposed. The inductance encountered by a small alternating-current ripple

[7] C. R. HANNA, *Trans. A.I.E.E.*, vol. 46, p. 155, February 1927. For the design of such choke coils see "Radio Engineering Handbook," McGraw-Hill Book Company, Inc., 2d ed., p. 476, 1935.

superimposed on larger values of direct current will be given by

$$L = \frac{0.4\pi N^2 A \mu_\Delta}{l \times 10^8} \text{ henrys} \tag{17}$$

where N is the total number of turns, A the area, and l is the mean length of the magnetic circuit, all dimensions being in centimeters. The inductance offered to the direct-current component can be obtained by substituting μ for μ_Δ. In many cases the direct current flowing through the coil is sufficient to produce a fairly high degree of saturation in the core material, resulting in a relatively small value of μ_Δ. If the reluctance of the magnetic circuit is increased by the insertion of a small air gap, as shown in Fig. 54, the resultant increase of μ_Δ caused by the reduced direct-current flux density may result in an increase in the alternating-current inductance, provided that the increase in μ_Δ is more than enough to offset the increased reluctance caused by the air gap. For any given conditions there will be a best length of air gap which will produce the maximum alternating-current inductance.

The total flux will be equal to the magnetomotive force divided by the total reluctance so that in the case of an iron core l centimeters in length containing an air gap of g centimeters, the direct-current flux will be

$$\Phi_{dc} = \frac{0.4\pi N I_{dc}}{\dfrac{l}{\mu A_1} + \dfrac{g}{A_2}} \tag{18}$$

where A_1 and A_2 are the respective areas of the core and air gap. The net area of the iron A_1 will be less than A_2, owing to the stacking factor caused by scale and air spaces between the laminations. The expression for the flux produced by the alternating current will be the same as (18) if μ_Δ is substituted for μ. The inductance in henrys offered to the alternating current will then be

$$L = \frac{N\Phi_{ac}}{I_{ac}} \times 10^{-8} = \frac{0.4\pi N^2}{\dfrac{l}{\mu_\Delta A_1} + \dfrac{g}{A_2}} \times 10^{-8} \tag{19}$$

It is often necessary to compute the inductance of an existing choke coil or transformer, having an air gap in the magnetic

circuit, when it is carrying a specified amount of direct current. The flux density in the core is not directly obtainable from magnetization curves of the material because of the fact that only a part of the total ampere-turns are consumed in overcoming the reluctance of the iron. But the reluctance of the iron depends on the permeability, which is not known until the flux density is known, since an empirical relation exists between μ and B. It is then necessary to plot a magnetization curve for the composite magnetic circuit of the choke coil in question by assigning various values of flux density and computing the necessary number of ampere-turns needed. The flux density corresponding to the given available number of ampere-turns can then be read from the curve. Knowing this value of flux density, μ or μ_Δ can then be obtained from the known data on the core material used. Only a few points in the vicinity of the actual flux density need be computed for the composite magnetization curve.

30. Condensers.—Electrostatic capacitance is present wherever a difference of potential can exist between two conducting bodies. The magnitude of the capacitance depends upon the area of the conductors, the distance between them, and also upon the nature of the dielectric traversed by the electrostatic lines of force. The capacitance can be calculated only when the conducting surfaces are comparatively simple geometric shapes. In the case where the conducting surfaces are parallel plates whose dimensions are large compared with the distance between them, so that the fringing of the electrostatic field at the edges may be neglected, the capacitance is given by

$$C = 0.2246K\frac{A}{d}\mu\mu\text{f} \qquad (20)$$

where A is the area of the active dielectric in square inches, d the distance between the plates in inches, and K is the dielectric constant of the material between the plates. The value of K is unity for air and from about 1.5 to 8 for the ordinary types of insulating materials.

Impregnated paper, mica, and air are the dielectrics most frequently used in condensers for radio work. Paper is relatively inexpensive but has the disadvantage of possessing appreciable losses due to dielectric hysteresis when an alternating voltage is impressed. These losses are similar in character to magnetic

hysteresis and are manifested in the form of heat generated within the dielectric. This energy loss causes the power factor of the condenser to be greater than zero. Air is free from dielectric hysteresis so that the power factor of air condensers is due chiefly to the losses in the necessary solid dielectric employed in their construction. There may also be a small amount of I^2R loss in the plates due to the charging current. The power factor of a condenser is substantially independent of the applied voltage and frequency, although it usually increases rather rapidly with increased temperature. In service, the rise in temperature due to dielectric losses brings about an increase in power factor which further increases the losses, so that the effect is cumulative. The dielectric strength of practically all insulating materials diminishes rapidly with an increase in temperature so that the condenser may break down in continuous service under voltages which could be withstood indefinitely in intermittent service. Wax-impregnated paper condensers undergo an aging process in service which weakens the dielectric after a period of months, so that a breakdown finally occurs. Thus, a group of condensers which successfully withstood a test of 600 volts direct current for 30 min. all broke down within a year of continuous service on 220 volts, 60 cycles. These characteristics of paper condensers make it difficult to devise acceptance tests or accelerated life tests from which definite conclusions may be drawn. Oil-impregnated paper is considered to be superior to wax-impregnated paper, particularly for high-voltage service.

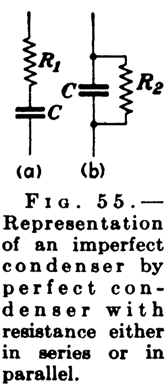


FIG. 55.—Representation of an imperfect condenser by perfect condenser with resistance either in series or in parallel.

The effect of the losses in a particular condenser may be taken into account by replacing the actual condenser with a perfect condenser of the same capacitance having a resistance in series as in Fig. 55*a*, or a resistance in parallel as in Fig. 55*b*. The equivalent series resistance in terms of the condenser power factor and capacitance may be determined as follows. The power factor is

$$\cos\theta = \frac{R_1}{\sqrt{R_1^2 + X_c^2}}$$

and since R_1 is small compared with X_c, we may write

$$R_1 = X_c \cos\theta = \frac{\cos\theta}{2\pi fC} \tag{21}$$

Likewise

$$R_2 = \frac{1}{2\pi fC \cos\theta} \tag{22}$$

31. Variable Condensers.—Variable condensers employing an air dielectric are used to adjust the resonant frequency of tuned circuits in radio receiving sets where a continuously variable capacitance having small losses is required. They are usually constructed of a series of fixed plates separated by washers, this group constituting one electrode called the stator. Interleaved between them with a small amount of clearance is a similar group of plates mounted on a rotatable shaft, which constitutes the movable electrode or rotor. The supporting frame, which is usually operated at ground potential, may be electrically connected to either one or the other of the electrodes giving rise to a *grounded-rotor* or *grounded-stator* construction. The capacitance of the condenser is determined by the angle of rotation of the shaft. By using semicircular rotor and stator plates the capacitance is approximately proportional to the angle of rotation so that condensers using this type of construction are called straight-line-capacity (SLC) condensers. This form of construction is rarely used in receivers designed for broadcast purposes since the various stations are separated by equal frequency intervals which would result in crowding the higher frequency stations together on the tuning dial of the condenser. By modifying the shape of either the stator or the rotor plates it is possible to have the capacitance proportional to the square of the angle of rotation. A condenser of this construction, when used with a given coil, causes the wave length at resonance to be proportional to the angle of rotation and is called a straight-line wave-length (SLW) condenser. The ideal condenser for broadcast reception is one designed so that resonant frequency is proportional to the angle of rotation. This type is termed a straight-line-frequency

(SLF) condenser.[8] It has the disadvantage of requiring a long, scimitar-shaped rotor plate pivoted near one end which requires more space for rotation than the other types. Both of these design types are affected by the distributed capacitance of the associated coil, and in order to get absolute straight-line wave-length or straight-line-frequency characteristics, it is necessary to design the coil to match a particular condenser, or vice versa. Most broadcast receiving sets use a tuning condenser lying somewhere between the straight-line wave-length and straight-line-frequency types. In order to simplify tuning adjustments in modern receiving sets, all the condenser rotors are mounted on a single shaft.

The ratio of maximum to minimum capacitance of the variable condenser governs the tuning range of the circuit if the coil inductance is fixed. Thus, if a 3:1 frequency range is to be covered, a 9:1 change in capacitance must be provided. The ratio of maximum to minimum capacitance of the ordinary variable condenser will be from about 10:1 to about 25:1, the latter ratio applying to the larger sized condensers. The distributed capacitance of the coil and the input capacitance of the vacuum tube or other associated apparatus must be taken into account along with the minimum condenser capacitance in computing the tuning range of the circuit.

The better grade of variable condensers for laboratory purposes are designed to have constant losses, independent of changes in the rotor setting. It is found that in these types the quantity $R_1\omega C^2$ (see Fig. 55*a*) is practically constant for all frequencies or scale readings and may be used as a figure of merit for the condenser. If its value is known, the equivalent series resistance can be determined for any frequency or scale reading. The value of this figure of merit is in the vicinity of 0.06×10^{-12} for a good laboratory condenser. Thus, a 1000-$\mu\mu$f condenser would have a series resistance of about 10 ohms at 1000 cycles and about 0.01 ohm at 10^6 cycles. Consequently, the resistance of a fairly good variable condenser is usually negligible at radio frequencies in comparison to the resistance of the associated coil.

[8] For design information on these condenser types see *Proc. I.R.E.* vol. 14, p. 773, December, 1926; also "Radio Engineering Handbook," McGraw-Hill Book Company, Inc., 2d ed., p. 122, 1935.

32. Electrolytic Condensers.—If two aluminum plates are immersed in a suitable electrolyte, such as a solution of ammonium borate or sodium phosphate, and connected to a source of direct current, a thin film of aluminum oxide forms on the positive plate which will gradually insulate the plate from the solution. The thickness of the film depends upon how large a voltage is used in its formation, higher voltages resulting in thicker films. Owing to the extremely small thickness of the film a capacitance of from 0.1 to 1.5 μf per square inch of area is obtained, depending on the forming voltage. Very large values of capacitance can therefore be obtained with only a fraction of the space required by paper-dielectric condensers. The film gradually disintegrates if the impressed voltage is removed from the cell. It is again formed when the voltage is reapplied, accompanied by a large leakage current which soon drops to the normal value of about 200 microamperes per microfarad, for working voltages of about 450. Over a period of several hundred hours of continuous operation the leakage current will gradually drop to only a few microamperes per microfarad. If an electrolytic condenser is operated for a considerable period of time at a voltage appreciably lower than the forming voltage used, the thickness of the film decreases and the capacitance rises. The maintenance of the oxide film requires that the anode be held at a positive potential, so that these condensers cannot be used on alternating current. Their chief use is in filter circuits of rectifiers where their comparatively large losses are of no consequence.

The maximum continuous working potential of these condensers is about 450 volts. Voltages much in excess of this value will puncture the film. However, the condenser is self-healing and the film is restored upon reducing the voltage. Higher voltage operation is readily secured by connecting two or more cells in series. The operation under this condition is satisfactory as the normal leakage resistance tends to equalize the voltage drop across each cell. When paper-dielectric condensers are operated in series across high direct-current voltages, it is usually necessary to shunt each unit with a high resistance to insure an equal division of voltage across the condensers. Otherwise the voltage will divide in accordance with the insulation resistances of the condensers and, while high in value, may be appreciably different for the several units.

In recent years "dry" electrolytic condensers have been developed wherein the electrolyte is in the form of a paste between two rolled foils. In other forms of construction a layer of gauze or other absorbent material is saturated with the electrolyte and then rolled up between the two aluminum sheets. Glycerin is usually mixed with the electrolyte because of its ability to absorb moisture. Condensers of this type are extremely compact and are available in very large values of capacitance for low voltage use. One commercial unit rated at 6000 μf at a working potential of 15 volts has a volume of only 100 cu. in. Units for 450 volts will have about a microfarad per cubic inch of volume, or about 15 per cent of the space required by paper condensers for the same voltage.

The life of these dry condensers depends to a considerable extent upon the conditions of use.

Problems

1. A coil having a diameter of 1.25 in. is wound with 100 turns of No. 30 enamel wire. If there are 88 turns per inch, what is the inductance of the coil? Compare this result with the value obtained from the empirical expression of equation (6).

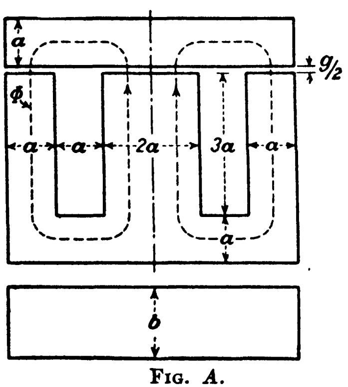


FIG. A.

2. The above coil was found to have a value of Q of 110 at 1000 kc. What is the ratio of the resistance at this frequency to the direct-current resistance, if No. 30 wire has a resistance of 103 ohms per 1000 ft.?

3. A multilayer coil of the form of Fig. 42 has a mean diameter of 0.75 in. The winding space is $\frac{3}{16}$ by $\frac{3}{16}$ in. How many turns are required to give an inductance of 15 mh, if the winding space is completely filled?

4. An audio-frequency transformer is wound on a shell-type core of 4 per cent silicon steel of the dimensions shown in Fig. A, where $g = 0$, $a = 0.375$ in., and $b = 0.875$ in. gross. Owing to scale and air spaces between the laminations, the stacking factor is 0.9, resulting in a net thickness of core of 0.7875 in. The windings are as follows: primary, 3000 turns of No. 40 enamel; secondary, 9000 turns of No. 42 enamel. From the magnetic data given in the table on page 75, determine the inductance of the primary for a small value of alternating current if the direct current carried by the primary is 0.005 amp. Assume that the mean path of the flux follows the center of gravity of the cross section as shown in Fig. A, and at the corners the path follows the quadrants of circles. What is the primary inductance if the direct current is reduced to 0.003 amp.?

MAGNETIC DATA

Flux density in lines per square centimeter	Type AW 4 per cent silicon steel		Hypernik 50 per cent nickel-iron	
B	μ	μ_Δ	μ	μ_Δ
2,000	5,200	1,248	47,800	10,800
3,000	6,450	1,242	51,400	9,950
4,000	7,500	1,235	49,600	8,400
5,000	8,300	1,195	45,500	6,880
6,000	8,840	1,130	40,300	5,560
7,000	9,050	1,035	34,500	4,450
8,000	8,800	920	28,500	3,530
9,000	8,200	790	23,000	2,800
10,000	7,350	655	17,600	2,190
11,000	6,160	520	13,000	1,700
12,000	4,740	380	9,600	1,400
13,000	2,950	240	7,220	1,100
14,000	1,550	120		
15,000	700	50		
16,000	300	20		

5. In Problem 4 what is the primary inductance offered to a direct current of 0.005 amp.? What voltage will be induced in the secondary if the flux produced by this current is reduced to zero at a uniform rate in 0.001 sec.?

6. If the transformer of Problem 4 is used as a step-down transformer so that the direct current through the 9000-turn winding is 0.005 amp., what will be the alternating-current inductance of that winding?

7. It is desired to raise the primary inductance of the transformer of Problem 4 by inserting an air gap g in the magnetic circuit so as to reduce the direct-current flux density in the iron to: (*a*) 8000, (*b*) 6000, (*c*) 4000 lines per square centimeter when the direct current in the primary is 0.005 amp. Find the required length of the gap in inches and the resulting alternating-current inductance in each case. Compare the results with Problem 4.

8. If the transformer of Problem 4 is replaced with a hypernik core of the same dimensions, determine the alternating-current inductance of the primary if the direct current in the winding is: (*a*) 0.005, (*b*) 0.003, (*c*) 0.001 amp.

9. An air gap is inserted in the magnetic circuit of Problem 8 so as to reduce the direct-current flux density in the iron to: (*a*) 8000, (*b*) 7000, (*c*) 6000 lines per square centimeter when the direct current in the primary is 0.005 amp. Find the required length of the gap in inches and the resulting inductance in each case.

10. A choke coil has 3000 turns of No. 27 enamel wire on a 4 per cent silicon steel, shell-type core of Fig. *A* having the following dimensions in

inches: $a = 0.625$, $b = 2$ (gross), $g/2 = 0.015$. The stacking factor is 0.9. What is the alternating-current inductance of the choke coil when carrying a direct current of 0.2 amp.? 0.1 amp.?

11. A wax-impregnated-paper condenser is to have a capacitance of 2 μf. How many square inches of dielectric must be provided if the thickness is 0.003 in. and the dielectric constant is 2?

12. An electrolytic condenser has a capacitance of 1.3 μf per square inch of anode. If the oxide film has an estimated thickness of 0.0001 mm., what is the dielectric constant of the film?

CHAPTER IV

COUPLED CIRCUITS

33. Mutual Inductance.—If two coils of N_1 and N_2 turns are so arranged with respect to each other that the magnetic field due to a current in one of them links in whole or in part with the other, a change in current in the first coil will induce an e.m.f. of mutual induction in the other.

This e.m.f. will be

$$e_2 = -N_2\frac{d\phi}{dt} \times 10^{-8} = -M\frac{di_1}{dt} \text{ volts} \tag{1}$$

where ϕ is the flux of the first coil that links with the second, and M is the mutual inductance between the two coils. From (1) we get

$$M = \frac{N_2\phi}{i_1} \times 10^{-8} \text{ henrys} \tag{2}$$

so that the mutual inductance is defined as *the number of flux linkages with one circuit per ampere of current in the other.* If the two coils have identical self-inductances L and are so related that all of the flux produced by the one coil links with all of the turns of the other, it is evident that the mutual inductance is equal to the self-inductance of either coil. In the event that the self-inductances are unequal, but with the same perfect degree of magnetic coupling, the mutual inductance is the geometric mean of the two self-inductances and will be

$$M = \sqrt{L_1L_2} \tag{3}$$

The coupling between two coils is never perfect in practice, so that the *coefficient of coupling* between two coils may be defined as

$$k = \frac{M}{\sqrt{L_1L_2}} \tag{4}$$

Consider the two coils connected series aiding in Fig. 56 with a source of potential e impressed across them. Applying Kirchhoff's laws to the circuit, we have at any instant

$$e = i_1R_1 + L_1\frac{di_1}{dt} + M\frac{di_2}{dt} + i_2R_2 + L_2\frac{di_2}{dt} + M\frac{di_1}{dt} \tag{5}$$

Since $i_1 = i_2 = i$,

$$e = i(R_1 + R_2) + (L_1 + L_2 + 2M)\frac{di}{dt} \tag{6}$$

If the two coils had been connected so that their magnetic fields were in opposition, the e.m.fs. of mutual induction in the coils would have been in opposite directions and the sign preceding $2M$ in (6) would have been negative. The total inductance of two coils connected in series with mutual inductance between them will therefore be

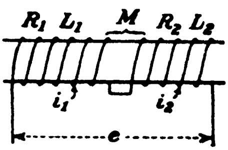


FIG. 56.—Two coils connected series aiding.

$$L_t = L_1 + L_2 \pm 2M \tag{7}$$

depending on whether they are connected series aiding or series opposing. Radio variometers, which are constructed so as to permit one coil to rotate within another through 180 degrees, make use of this principle. The total variation in inductance obtained in this manner will be $4M$.

If the impressed voltage in Fig. 56 had been a sine wave, the steady-state solution of the differential equation (6) would have been

$$E = I\sqrt{(R_1 + R_2)^2 + \omega^2(L_1 + L_2 \pm 2M)^2} \tag{8}$$

or in vector form,

$$E = I(R_1 + R_2) + jI\omega(L_1 + L_2 \pm 2M) \tag{9}$$

When the two coils are not conductively connected to each other, as in the case of the primary and secondary of a transformer, the voltage induced in the secondary is always in opposition to the agent that produced it. Thus, the open-circuit voltage E_2 induced in the secondary coil due to a sinusoidal current I_1 in the primary coil is

$$E_2 = -jI_1\omega M \tag{10}$$

This expression holds regardless of how the secondary coil may be wound with respect to the primary.

34. Transformers.—Any two coils having mutual inductance between them can serve as a transformer. For the lower range of communication frequencies the two coils are wound on a magnetic core to insure a high degree of coupling between them. At high frequencies the loss in the core becomes a serious factor, as discussed in the preceding chapter, and an air core is used.

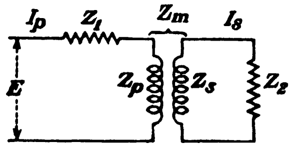


FIG. 57.—Transformer with an impedance Z_1 in series with the primary and a load Z_2 connected across the secondary.

The circuit of Fig. 57 represents a typical transformer, either air core or iron core, with a load Z_2 connected across the secondary and an impedance Z_1 in series with the primary. These impedances may be resistance, reactance, or any combination of these elements. The impedance looking into the primary with the secondary on open circuit is Z_p, and Z_s is the impedance of the secondary with the primary open. The mutual impedance Z_m will be of the nature of a pure reactance $j\omega M$ in the case of an air-core transformer. With an iron core Z_m will contain, in addition to the mutual-reactance term a resistance component which simulates the core loss. Applying Kirchhoff's laws to the circuit, we get

$$E = I_pZ_1 + I_pZ_p + I_sZ_m \tag{11}$$

$$0 = I_sZ_2 + I_sZ_s + I_pZ_m \tag{12}$$

From (12)

$$I_s = -I_p\frac{Z_m}{Z_2 + Z_s} \tag{13}$$

Substituting this value of I_s in (11),

$$E = I_p(Z_1 + Z_p) - I_p\frac{Z_m^2}{Z_2 + Z_s}$$

or

$$I_p = \frac{E(Z_2 + Z_s)}{(Z_1 + Z_p)(Z_2 + Z_s) - Z_m^2} \tag{14}$$

Likewise

$$I_s = \frac{-EZ_m}{(Z_1 + Z_p)(Z_2 + Z_s) - Z_m^2} \tag{15}$$

It is to be understood that the above are vector equations and the complex expressions for the various impedances must be

used. The negative sign of (15) indicates that the secondary current is 180 degrees out of phase with the primary current. The impedance looking into the primary circuit from the source of the applied voltage E is

$$Z = \frac{E}{I_p} = \frac{(Z_1 + Z_p)(Z_2 + Z_s) - Z_m^2}{Z_2 + Z_s} = Z_1 + Z_p - \frac{Z_m^2}{Z_2 + Z_s} \quad (16)$$

The term $\frac{Z_m^2}{Z_2 + Z_s}$ is the impedance reflected into the primary by the secondary. When using generalized impedances the sign preceding this term is negative, as shown. The denominator will usually be of the form $R \pm jX$, so that the reflected impedance will contain, in addition to a resistive term of positive sign, a reactive term whose sign will depend upon the nature of the total reactance in the secondary circuit. The sign of this reflected reactance is always opposite to that of the total secondary reactance. In other words, if $Z_2 + Z_s$ is inductive, the reflected term will be of the nature of $-jX$, which will reduce the inductive reactance looking into the primary. With perfect coupling and Z_2 a pure resistance, this reflected term will be just large enough to annul completely the inductive reactance of the primary, and the impedance looking into the primary will be of the nature of a pure resistance.

The foregoing equations are applicable to any type of transformer and will be frequently used. In power transformers the coefficient of coupling is nearly unity, resulting in a ratio of transformation that is very nearly equal to the turn ratio. Computations involving these transformers can be more readily handled by dealing with the reactances caused by the small amount of leakage flux present. In radio circuits, particularly where the primary and secondary are tuned to resonance, the coefficient of coupling may be very small and consequently the leakage reactances are often 95 per cent or more, so that the effective ratio of transformation is quite different from the ratio of turns on the transformer. As a result of these differences it is much easier to analyze the coupled circuits used in radio work in terms of mutual inductance rather than by turn ratio and leakage reactance.

35. Coupled Resonant Circuits.—One common type of coupled circuit used in radio consists of an antenna circuit inductively

coupled to a tuned secondary circuit, as shown in Fig. 58, where *a* is the actual circuit and *b* is the electrical equivalent. We are usually interested in the conditions which will make the secondary current I_2 a maximum so as to develop maximum voltage across C_2. Circuits of this kind are usually followed by an amplifier, which is a voltage-operated device. The input capacitance of the amplifying tube may be included in the value

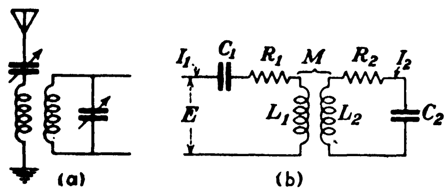


FIG. 58.—Tuned secondary coupled to an antenna circuit

of C_2. The impressed voltage E will be the total voltage induced in the antenna and will be

$$E = he \tag{17}$$

where h is the effective height[1] of the antenna in meters and e is the field strength of the received signal expressed in volts (or microvolts) per meter.

The current in the secondary is given by (15) where

$$Z_1 + Z_p = R_1 + j\left(\omega L_1 - \frac{1}{\omega C_1}\right) \equiv R_1 + jX_1$$

$$Z_2 + Z_s = R_2 + j\left(\omega L_2 - \frac{1}{\omega C_2}\right) \equiv R_2 + jX_2$$

$$Z_m = j\omega M$$

Substituting these values in (15),

$$I_2 = \frac{-Ej\omega M}{(R_1 + jX_1)(R_2 + jX_2) + \omega^2 M^2} = \frac{-Ej\omega M}{(R_1R_2 - X_1X_2 + \omega^2M^2) + j(X_1R_2 + X_2R_1)} \tag{18}$$

or in absolute magnitude

$$|I_2| = \frac{E\omega M}{\sqrt{(R_1R_2 - X_1X_2 + \omega^2M^2)^2 + (X_1R_2 + X_2R_1)^2}} \tag{19}$$

[1] The effective height h does not bear any simple relation to the actual antenna height other than that defined by (17).

The secondary current I_2 will be a maximum when the denominator of (19) is a minimum, or for convenience in differentiating, when the square of the denominator is a minimum. If X_2 is considered to be the variable, I_2 will be a maximum when

$$X_2 = X_1 \frac{\omega^2 M^2}{R_1^2 + X_1^2} \tag{20}$$

In other words, if the antenna circuit is not resonant to the frequency of the impressed signal (X_1 not equal to zero), the secondary must be detuned from resonance by the amount indicated by (20). If the secondary circuit is fixed in adjustment and the antenna circuit is tunable, we find upon differentiating (19) with respect to X_1 that I_2 will be a maximum when

$$X_1 = X_2 \frac{\omega^2 M^2}{R_2^2 + X_2^2} \tag{21}$$

If both primary and secondary circuits are simultaneously adjustable, I_2 will be a maximum when X_1 and X_2 are both made zero, provided ωM is not greater than the critical value which will be defined later. The expression for the secondary current will then be

$$I_{2\text{max}} = \frac{E\omega M}{R_1 R_2 + \omega^2 M^2} \tag{22}$$

If the coupling between L_1 and L_2 is now varied, keeping X_1 and X_2 equal to zero, differentiating (22) with respect to M shows that maximum secondary current will be had when

$$\omega M = \sqrt{R_1 R_2} \tag{23}$$

Substituting this value of ωM in (22) we get

$$I_{2\text{opt}} = \frac{E}{2\sqrt{R_1 R_2}} \tag{24}$$

This represents the optimum condition and is the maximum possible value of I_2 that can be obtained from a source of e.m.f. by any passive network such as a transformer.

These relations can be more readily visualized by the following treatment due to G. W. Pierce.[2]

[2] "Electric Oscillations and Electric Waves," Chap. XII, McGraw-Hill Book Company. Inc., 1920.

Substituting the value of X_2 in (20) we have

$$\omega L_2 - \frac{1}{\omega C_2} = \frac{\left(\omega L_1 - \frac{1}{\omega C_1}\right)\omega^2 M^2}{R_1^2 + \left(\omega L_1 - \frac{1}{\omega C_1}\right)^2}$$

or

$$\omega L_2\left(1 - \frac{1}{\omega^2 L_2 C_2}\right)\left[R_1^2 + \omega^2 L_1^2\left(1 - \frac{1}{\omega^2 L_1 C_1}\right)^2\right] = \omega L_1\left(1 - \frac{1}{\omega^2 L_1 C_1}\right)\omega^2 M^2$$

Defining

$$\Omega_1^2 = \frac{1}{L_1 C_1}, \quad \Omega_2^2 = \frac{1}{L_2 C_2}, \quad k^2 = \frac{M^2}{L_1 L_2}, \quad Q_1 = \frac{\omega L_1}{R_1}, \quad Q_2 = \frac{\omega L_2}{R_2},$$

and substituting them in the above expression,

$$\frac{1}{Q_1^2}\left(1 - \frac{\Omega_2^2}{\omega^2}\right) + \left(1 - \frac{\Omega_2^2}{\omega^2}\right)\left(1 - \frac{\Omega_1^2}{\omega^2}\right)^2 = k^2\left(1 - \frac{\Omega_1^2}{\omega^2}\right)$$

or

$$\left(1 - \frac{f_2^2}{f^2}\right)\left(1 - \frac{f_1^2}{f^2}\right) = k^2 - \frac{1}{Q_1^2}\frac{1 - \frac{f_2^2}{f^2}}{1 - \frac{f_1^2}{f^2}} \tag{25}$$

where f is the impressed frequency and f_1 and f_2 are the respective resonant frequencies of the primary and secondary circuits alone.

If the secondary tuning had been fixed and the primary had then been adjusted for maximum current in the secondary, an expression would have been obtained from (21) which would have been identical in form with (25), except that all subscripts would have been interchanged, 1 becoming 2 and vice versa.

Equation (25) represents the locus of all maxima of secondary current and is plotted in Fig. 59 for $k = 0.5$ and various values of Q_1. For example, if the antenna circuit has a resonant frequency of 1800 kc, and a signal voltage of 1000 kc was induced in it, f_1/f would be 1.8, and from Fig. 59 the corresponding value of f_2/f is found to be 1.054. The secondary circuit would, therefore, have to be tuned to 1054 kc for maximum secondary current if the coefficient of coupling is 0.5.

For finite values of Q_1 the curve approximates a hyperbola for large values of f_1/f. As this ratio approaches unity, the curve sweeps down through the point (1, 1) and into the third quadrant. The curve $Q_1 = 10$ has a gap in it in the interval from 0.9745 to 0.894, which merely means that there is no setting of the tuning

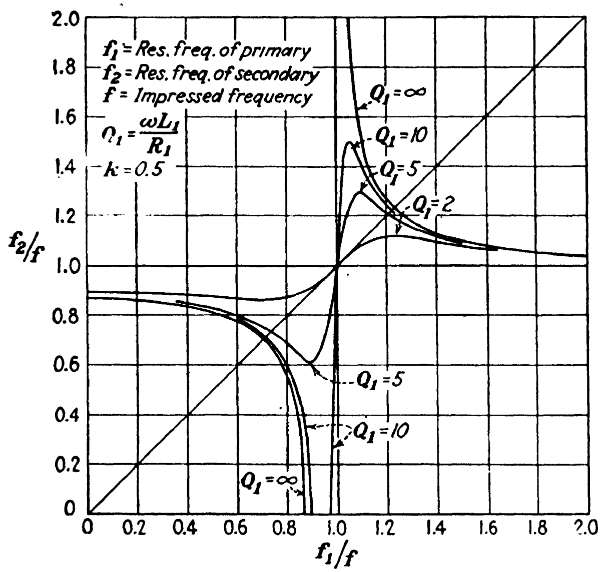


FIG. 59.—Locus of maximum secondary current for various values of primary reactance expressed as a ratio of the resonant frequency f_1 of the primary to the impressed frequency f. With this ratio known the corresponding adjustment f_2/f of the secondary is given by the curves for a coefficient of coupling of 0.5 and various values of Q_1.

condenser C_2 which will produce a mathematical maximum of I_2 between these values of f_1/f.

If the resistance in the primary circuit is small enough so that $1/Q_1^2$ is negligible, equation (25) becomes

$$\left(1 - \frac{f_2^2}{f^2}\right)\left(1 - \frac{f_1^2}{f^2}\right) = k^2 \tag{26}$$

which is the equation of an equilateral hyperbola in terms of $(f_2/f)^2$ versus $(f_1/f)^2$ with asymptotes at $f_1/f = 1 = f_2/f$. A family of such curves is plotted in Fig. 60 for various values of k. Since the first powers of the frequency ratios are used instead of the squared values, the curves are not symmetrical about the axes (1, 1). For values of Q_1 greater than 50 there is no appreciable

difference between the results of (25) and (26) except for very small values of k. As the coefficient of coupling is made smaller, the two hyperbolas approach the axes (1, 1) and coincide with them when $k = 0$. If the primary and secondary are both tuned to the same frequency and the impressed frequency is then varied, the values of f_2/f and f_1/f will move along a 45-degree line through the point (1, 1). As will be noted in Fig. 60, the secondary current will have two maximum values at the inter-

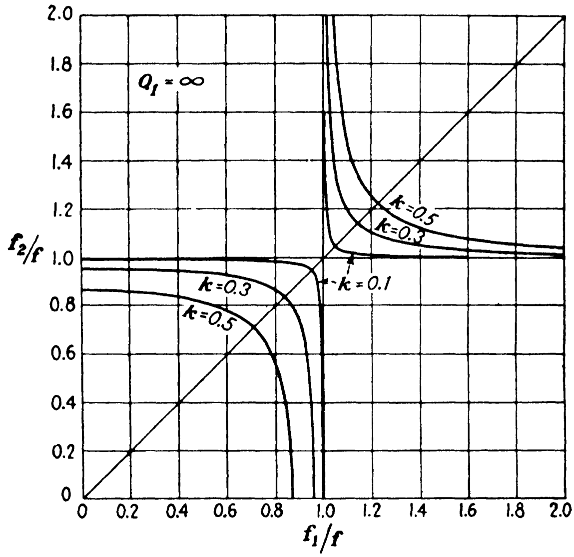


FIG. 60.—Locus curves of maximum secondary current for various coupling coefficients.

section of the 45-degree line with the pair of hyperbolas. As the coupling is diminished, the two maxima approach each other, theoretically merging into a single value only when $k = 0$. In actual practice the circuit resistance is never zero so that the two peaks of secondary current merge into a single maximum when $\omega M = \sqrt{R_1R_2}$, this condition being termed *critical coupling.* Substituting Q_1, Q_2, and k, this relation becomes

$$\omega k\sqrt{L_1L_2} = \sqrt{\frac{\omega^2L_1L_2}{Q_1Q_2}}$$

$$k = \frac{1}{\sqrt{Q_1Q_2}} \tag{27}$$

If $Q_1 = Q_2 = Q$, critical coupling becomes

$$k = \frac{1}{Q} \tag{28}$$

The family of curves in Fig. 60 may be viewed as the map of a group of mountain ranges whose elevation at any point corresponds to the value of secondary current. These ridges have their maximum elevation at the intersection with the 45-degree line and taper off in height as the two ends of the hyperbolas are approached. A cross section along the 45-degree line for a typical case is shown in Fig. 61 for various values of coupling above and below critical value.

For convenience, reciprocals of the previously used frequency ratios are plotted in the upper half of the figure, since the impressed frequency f is now considered as the independent variable. The general shape of the locus curves will be much the same as in Fig. 59. The family of curves has been rotated so that the 45-degree line is now horizontal.

It will be observed that as the coupling is reduced, the two peaks of secondary current approach each other and merge into a single value at critical coupling. The top of the resonance curve at this value of coupling is flatter and the sides are steeper than the corresponding series or parallel resonance curves of Figs. 21 and 35. Coupled circuits using couplings equal to, or slightly in excess of, critical value are often used where a narrow band of frequencies is to be transmitted with fairly uniform response, as in broadcast reception. Circuits of this type are termed band-pass filters. The width of the band that can be uniformly transmitted is limited by the increasing depth of the hollow between humps as the coupling is increased. A smaller value of Q tends to reduce the depth of the hollow, but it also reduces the discrimination against unwanted frequencies lying without the transmitted band. A large value of Q produces very pronounced double humps.

The approximate position of these two resonant peaks can be determined by solving for f in (26) which gives

$$f = \sqrt{\frac{f_1^2 + f_2^2 \pm \sqrt{(f_1^2 - f_2^2)^2 + 4k^2f_1^2f_2^2}}{2(1 - k^2)}} \tag{29}$$

The plus sign under the radical gives the high-frequency peak

and the negative sign the lower. If both circuits are tuned to the same frequency so that $f_1 = f_2 = f_0$, (29) becomes

$$f = \frac{f_0}{\sqrt{1 \pm k}} \tag{30}$$

This expression becomes inaccurate as k approaches critical coupling, owing to the fact that Q has been assumed to be infinite

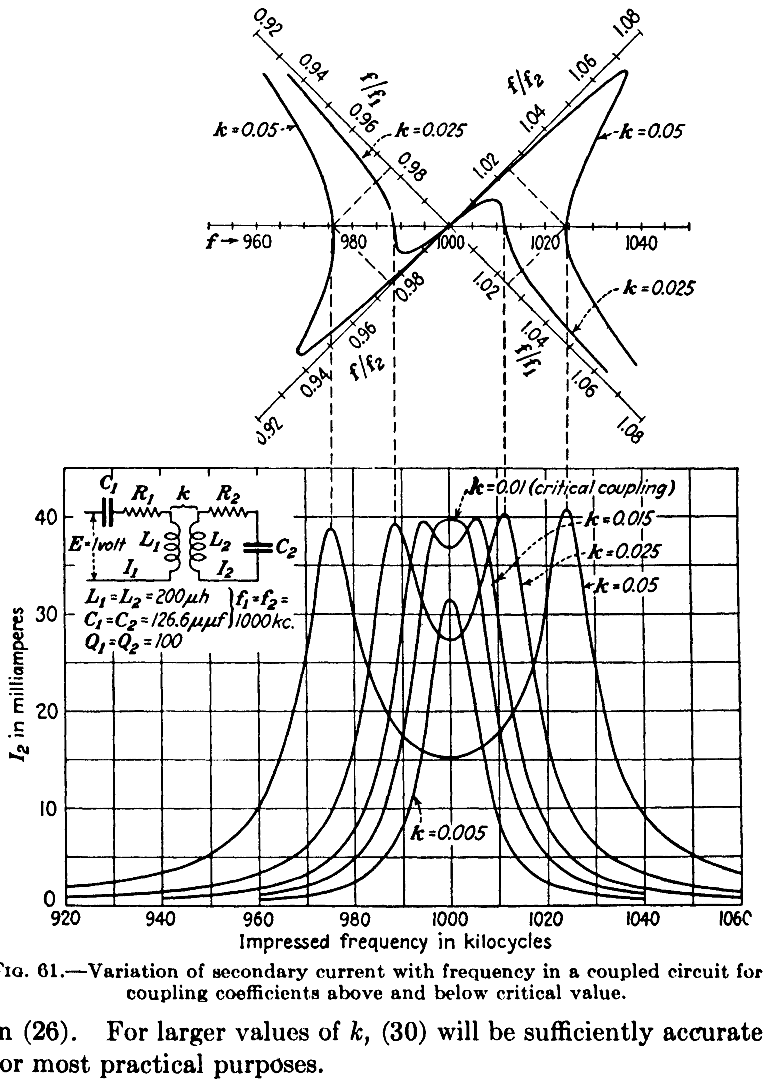


FIG. 61.—Variation of secondary current with frequency in a coupled circuit for coupling coefficients above and below critical value.

in (26). For larger values of k, (30) will be sufficiently accurate for most practical purposes.

A reduction of k below critical value causes a reduction in the maximum value of I_2 and a more sharply defined peak. It will be noted that the maximum value of I_2 is substantially constant for all values of coupling equal to, or in excess of, critical. If the resistance of the circuit instead of Q had remained constant with frequency, these peaks would have all been of the same height. With Q constant, which approximates practical conditions for small ranges of frequency, the high-frequency peak becomes taller and the low-frequency peak becomes lower as k is increased.

At critical coupling there is only one intersection of the 45-degree line with the curve representing the locus of maximum secondary current. The curve $Q_1 = 2$ in Fig. 59 is typical of the locus curve at critical coupling.

Viewing the family of locus curves as the ridges of mountain ranges is also useful in understanding the differences observed in the sharpness of tuning of the primary and secondary circuits at various points. This is illustrated in Fig. 62 where 1 volt at a frequency of 1000 kc is impressed on the primary of the circuit shown. The setting of condenser C_1 is such as to resonate the primary circuit to a frequency of 1500 kc. If C_2 is then varied, I_2 will be a maximum at point A when f_2 is 1015.86 kc. As f_2 is varied, we are crossing the locus of maximum secondary current along the line aa'. This line crosses the ridge at nearly a right angle so that the secondary tuning is sharp as shown by the resonance curve drawn to the right of point A. If C_2 is left at the point of maximum I_2 and f_1 is now varied by means of C_1, the variation will be along the line bb'. Since this line is nearly parallel to the ridge, the primary tuning is broad, as shown by the resonance curve above the point A. Furthermore, I_2 will now be a maximum when f_1 is set for about 1465 kc as the elevation of the ridge rises considerably as we approach B, the point of maximum elevation for this value of coupling. The mathematical explanation is that the locus is derived on the basis of f_2 as the independent variable; consequently the coordinates of point A will result in maximum secondary current only when f_2 is the variable.

At point B the sharpness of tuning will be practically the same for either circuit, the secondary tuning having become somewhat broader than at A. If the primary circuit is resonated to a

frequency only slightly above the impressed frequency, as at point C, we find the conditions just reversed from what they were at point A; the primary tuning being now quite sharp while the secondary tuning is extremely broad. A similar condition exists in the third quadrant of the locus curves. Here the frequency of the primary circuit will lie *below* the impressed

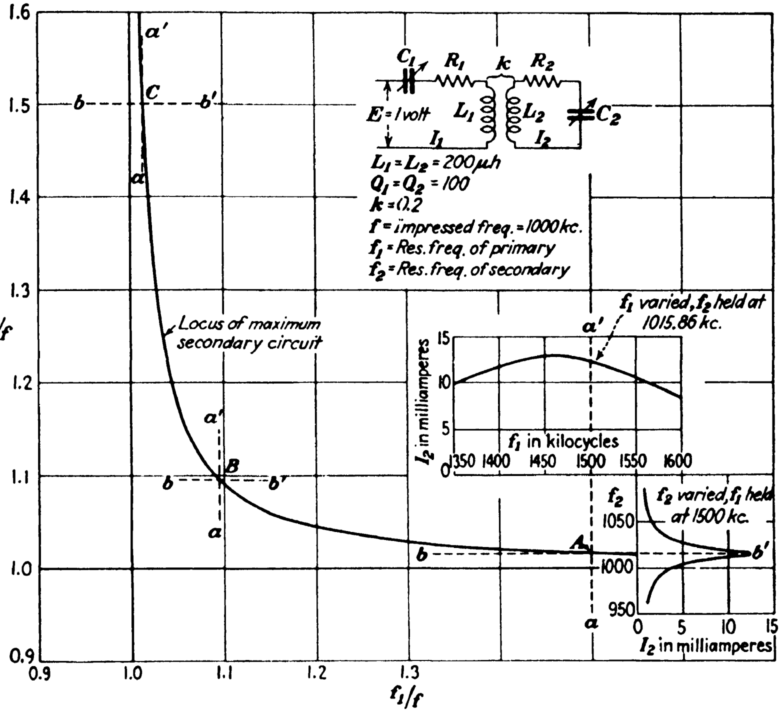


FIG. 62.—Comparison of the sharpness of tuning of the primary and secondary circuits when the primary is tuned to a frequency 50 per cent higher than the impressed frequency.

frequency. An identical shift in the sharpness of tuning from one circuit to the other occurs in this region. In general, the secondary tuning will be broad whenever the primary constants are such that f_1/f is near unity. Under this condition the primary tuning is quite sharp. As the ratio f_1/f recedes from unity the secondary tuning becomes progressively sharper, while the primary tuning becomes broader.

These locus curves may be used also to predict the behavior of coupled circuits when the primary and secondary are tuned to

different frequencies and the impressed frequency is varied, by drawing a line through the origin having the proper slope. For example, suppose f_1 is 1000 kc and f_2 is 1100 kc. Then in either Fig. 59 or 60 a line drawn from the origin passing through the point $f_2/f = 1.1$, $f_1/f = 1$, will represent the variation in the impressed frequency. The intersection of this line with the pair of hyperbolas corresponding to the value of k of the circuit will locate the resonance peaks. An idea of the relative heights of the two peaks may be gained by noting the location of the intersection as compared with the intersection of the 45-degree line, where the secondary current is a maximum. The shapes of the curves can also be roughly estimated, following considerations similar to those used in connection with Fig. 62.

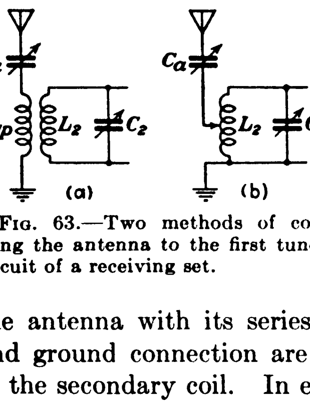


FIG. 63.—Two methods of coupling the antenna to the first tuned circuit of a receiving set.

36. Antenna-circuit Adjustments in Receiving Sets.—A large number of receiving sets for broadcast reception have the first tuned circuit coupled to the antenna in the manner of Fig. 63*a* or else use an autotransformer as in Fig. 63*b*. In the latter case the antenna with its series condenser C_a (sometimes omitted) and ground connection are connected across the lower portion of the secondary coil. In either case the coupled circuit theory of the preceding section is applicable. The constants of the antenna circuit are reflected into the secondary so that the setting of the secondary tuning condenser for maximum secondary current will be somewhat different than for the condensers of the subsequent tuned stages. These condensers are all ganged together in modern sets so that it is necessary to prevent the antenna constants from detuning the first stage. The value of capacitance required by the first tuned circuit is less than for the other stages, since an antenna circuit operated at a frequency below the resonant frequency reflects *inductance* into the secondary coil. This is taken care of by shunting the other tuning condensers by a small "trimming" condenser. These trimming condensers also compensate for variations in the capacitance shunted across the various tuned stages caused by differences in lengths of lead wires, tube-input capacitance, etc. If the ratio

of f_1/f for the antenna circuit is large, as would be the case with a small antenna, we shall be operating on the nearly horizontal portion of the hyperbola of Fig. 60, and the reactance reflected into the secondary will be nearly constant over the entire tuning range of the circuit. This is merely another way of stating that f_2/f is practically constant. However, if the coefficient of coupling is large or the natural frequency f_1 of the antenna circuit is too low, the value of f_2/f will change appreciably over the tuning range. This will prevent the several tuned circuits from remaining in tuning alignment by a fixed setting of the trimming condensers. In order to avoid this difficulty, the antenna circuit should have a resonant frequency which is considerably higher than the highest frequency to which the set can be tuned. This is usually accomplished by making the maximum capacitance of the adjustable condenser C_a rather small. It is usually adjusted by means of a screw driver when the set is connected to the antenna. The purpose of the adjustment is to reflect the proper amount of reactance into the secondary circuit so that the first tuning condenser will be properly aligned with the others, and is not to adjust the antenna circuit to resonance. From the discussion relative to Fig. 62, this tuning adjustment will be rather broad. Some sacrifice in sensitivity is necessarily made by operating out on this portion of the hyperbola, as the maximum value of secondary current is much smaller here than the values obtainable nearer to the intersection of the 45-degree line. This is not serious in modern sets which readily obtain by subsequent amplification all the sensitivity that can be usefully employed.

It is also possible to prevent the constants of the antenna circuit from affecting the alignment of the tuning condensers by going to the other extreme and making the resonant frequency of the antenna circuit lower than the lowest frequency to which the set can be tuned. Receiving sets employing this principle use a very large inductance L_p in the antenna circuit which causes it to have a value of f_1 well below the lowest frequency we desire to receive. A similar result could have been obtained by the use of an antenna large enough to have had natural frequency in this same region, in which case only a nominal value of L_p would be required. The antenna series condenser is not needed as it tends to raise the frequency of the antenna circuit, which is not

desired. Since all the frequencies to be received will be higher than the natural period of the antenna circuit, *negative reactance* will be reflected into the secondary coil and the effective value of L_2 will be reduced. This requires more tuning capacitance across L_2 than is required by the subsequent stages, which is just the opposite to the previous case. A trimming condenser shunted across C_2 in Fig. 63*a* can then be used to compensate for the reflected constants of the antenna used. In this case the use of too large an antenna will not produce the alignment difficulties caused by the failure of the first tuning condenser to properly track with the others. Possible difficulties with too small an antenna can be avoided by using a primary coil L_p of sufficient size so that its inductance, in conjunction with its distributed capacitance, produces a low enough value of f_1 without an antenna. Then any additional capacitance shunted across L_p in the form of a small antenna merely brings about a further reduction in f_1. Another advantage of operating with f_1/f less than unity is that the selectivity of the first stage tends to become more nearly uniform over the tuning range.

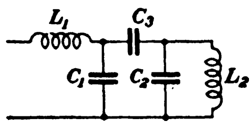


FIG. 64.—Capacitively coupled circuits.

37. Other Forms of Coupling.—In addition to inductive coupling, a transfer of energy from one circuit to another may be accomplished by capacitive coupling as in Fig. 64, where C_3 is the coupling condenser. Circuits of this type have characteristics similar in all respects to the inductively coupled circuits just considered. Reducing the size of C_3 reduces the coefficient of coupling, which in this case is defined as

$$k = \frac{C_3}{\sqrt{(C_1 + C_3)(C_2 + C_3)}} \tag{31}$$

Combinations of inductive and capacitive coupling are illustrated in Fig. 65. The mutual inductance can be made to induce a voltage in the secondary circuit which either aids or opposes the voltage due to capacitive coupling, depending on the sign of M. The first voltage varies directly with the frequency while the second varies inversely, making it possible to secure a coefficient of coupling that varies with the frequency. This is of use when it is desired to transmit a band of frequencies of constant width and to allow this band to be shifted over the broadcast

range by tuning, without substantial change in its width. Figure 65*b* shows a circuit of this type described by E. A. Uehling.[3] A constant band width over an appreciable tuning range cannot be secured by means of magnetic coupling alone unless some

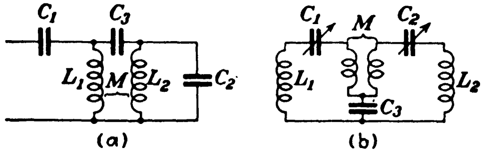


FIG. 65.—Circuits with both capacitive and inductive coupling.

mechanical means is employed to vary the coupling with the tuning.

Problems

1. A radio-frequency transformer has a secondary inductance of 240 μh. The primary and secondary are connected in series and the total inductance is found to be 349.5 μh. With the coils in series but with the connections to the primary reversed, the total inductance under this condition is 229.5 μh. What is the primary inductance? What are the mutual inductance and the coefficient of coupling?

2. Two coils L_1 and L_2 having an inductance of 0.01 and 0.09 henry, respectively, have a coefficient of coupling between them of 0.8. If 1 volt at 10^5 cycles is impressed across L_1, what will be the open-circuit voltage across L_2? Assume that the resistance of L_1 is negligible compared with its reactance.

3. In the circuit shown in Fig. *A*, $1/\omega C_2$ is made equal to ωL_2. The condenser C_1 is then adjusted so that the impedance looking into the terminals *a* and *b* is a pure resistance. What is the magnitude of this impedance at 10^6 cycles if $R_1 = R_2 = 10$ ohms and $M = 10\ \mu$h?

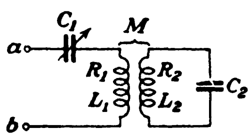


FIG. *A*.

4. In Problem 3 what would be the voltage across C_2 if the potential impressed across *a* and *b* is 10 volts at 10^6 cycles? What would be the secondary voltage if *M* were adjusted to its optimum value? What is the optimum value of *M*? $C_2 = 200\ \mu\mu f$.

5. A coupled circuit similar to Fig. *A*, except that C_1 is removed, has the following constants: $R_1 = 10$ ohms, $L_1 = 100\ \mu$h, $R_2 = 20$ ohms, $L_2 = 200\ \mu$h, $M = 25.82\ \mu$h. What must be the value of C_2 so that the impedance looking into the primary shall be a pure resistance, if $\omega = 6 \times 10^6$? What is the magnitude of this resistance?

6. The total constants of an antenna circuit similar to Fig. 58 are as follows: $L_1 = 121.5\ \mu$h, $C_1 = 200\ \mu\mu$f, $R_1 = 25$ ohms. Coupled to it with a mutual inductance of 20 μh is a secondary circuit of $L_2 = 200\ \mu$h and $R_2 = 10$ ohms. Find the value of C_2 in order that the secondary current

[3] *Electronics*, vol. 1, p. 279, September, 1930.

shall be a maximum if the frequency of the impressed signal is 10^6 cycles. How does this value of C_2 compare with the value that would be needed to resonate with L_2 if the antenna circuit were absent?

7. In Problem 6 the effective height of the antenna is 4 meters and the intensity of the received signal is 10 millivolts per meter. What is the voltage across C_2 when the latter is adjusted for maximum received current?

8. The antenna circuit of a short-wave receiver has a resonant frequency of 2000 kc. The secondary circuit is coupled to the antenna with a coefficient of coupling of 0.2. To what frequency must the secondary be tuned in order that the secondary current shall be a maximum for a signal frequency of 6000 kc? The effective Q of the antenna circuit is 10 at this frequency.

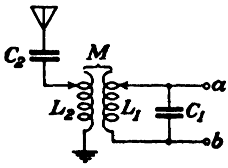


FIG. B.

9. The antenna coupling system of a transmitting station is shown in Fig. B and it is desired to make the impedance looking into a and b equal to 1000 ohms, pure resistance, by proper adjustment of L_1 and M. The antenna has a natural frequency of 10^6 cycles and a resistance of 50 ohms with L_2 and C_2 both removed. Since it is desired to operate at 10^6 cycles, a condenser C_2 of 0.0002 μf is inserted in series with the antenna and L_2 is adjusted to resonate the antenna circuit to this frequency. The tank circuit L_1C_1 is then coupled to the antenna circuit and adjusted, C_1 being 0.002 μf. The apparent Q looking into the terminals of coil L_1 is constant as M and L_1 are adjusted. The actual resistances of L_1 and L_2 are negligible. Find the necessary values of L_1, L_2, and M.

10. If 1000 volts at 10^6 cycles is impressed across a and b in Problem 9, find the antenna current and the currents in L_1 and C_1.

11. An output transformer has the following constants at 400 cycles: $Z_p = 4850 + j65{,}000$ ohms, $Z_s = 6.5 + j88$ ohms. With Z_p short-circuited the impedance looking into Z_s is $3.91 + j2.4$ ohms. What is Z_m?

12. A load of 10 ohms resistance is connected across the secondary of the above transformer. If 100 volts at 400 cycles is impressed across the primary what will be the current in the load? What will be the primary current? What will be the power input and output?

13. Show that in the case of a coupled circuit having identical primary and secondary constants, as in Fig. 61, the expression for the secondary current is given by

$$I_2 = \frac{EkQ^2}{2\pi fL\left[1 + k^2Q^2 - Q^2\left(1 - \frac{f_0^2}{f^2}\right)^2 + j2Q\left(1 - \frac{f_0^2}{f^2}\right)\right]}$$

where f_0 is the natural frequency of the primary and secondary alone.

CHAPTER V

OSCILLATORY CIRCUITS

38. Free Oscillations and Mechanical Analogies.—The circuits discussed in the preceding chapters may be looked upon as cases of forced oscillation. An external voltage of fixed frequency was impressed on the circuit so that the steady-state current which resulted was always of a frequency dictated by the source. The problem is analogous to the case in mechanics of a pendulum or a mass supported by a spring which is acted upon by a periodic recurrent force. The fundamental period of the resultant motion will be the same as that of the applied force. The amplitude of motion, which is analogous to the current in the electrical case, will depend upon the magnitude of the force and its frequency relative to the natural frequency of the vibrating system. When the disturbing force is of the same frequency as the natural oscillatory period of the system a relatively large amplitude of vibration can be set up by means of an extremely small force. This is exactly analogous to the case of series resonance. Thus, if a resonant circuit is electrically disturbed it will oscillate at its resonant frequency. The frequency of this free oscillation, however, is affected to some extent by the resistance in the circuit.

The differential equations of the electrical circuit have their exact equivalent in mechanics; inductance being equivalent to mass, resistance to friction, and capacitance to compliance, the reciprocal of stiffness. Mechanical analogies have been extensively used to explain electrical-circuit behavior, but in recent years electrical-network theory and methods of analysis have progressed to such a degree that the complex problems of mechanical vibration, such as are met with in the design of various acoustic devices, are being handled by means of their equivalent electrical networks. This method of attack has been responsible for marked strides in the field of applied acoustics.

39. Charge of a Condenser through an Inductance and Resistance in Series.—If a source of potential, such as a battery, is

applied to a circuit composed of R, L, and C in series by closing the switch S in Fig. 66, the impressed voltage at any instant is given by

$$E = iR + L\frac{di}{dt} + \frac{q}{C} \tag{1}$$

where q is the charge on the condenser in coulombs. Since

$$q = \int idt$$

(1) becomes

$$E = iR + L\frac{di}{dt} + \frac{1}{C}\int idt$$

which upon differentiation is

$$\frac{dE}{dt} = 0 = R\frac{di}{dt} + L\frac{d^2i}{dt^2} + \frac{i}{C} \tag{2}$$

This is a linear differential equation of the first degree and of the second order (contains second derivatives), with constant coefficients. All linear differential equations have as their solutions an equation of the form

$$i = A\epsilon^{kt} \tag{3}$$

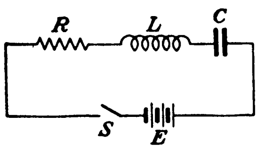


FIG. 66.—Circuit composed of R, L, and C in series connected to a source of potential by means of switch S.

where A and k are constants whose values are determined by the limiting physical conditions of the problem. Equation (3) is only the type form as there are always as many terms in the complete solution as there are units in the order of the equation—in this case, two. The complete solution may then be expected to be of the form

$$i = A_1\epsilon^{k_1t} + A_2\epsilon^{k_2t} \tag{4}$$

Differentiating (3) twice

$$\begin{aligned} \frac{di}{dt} &= Ak\epsilon^{kt} \\ \frac{d^2i}{dt^2} &= Ak^2\epsilon^{kt} \end{aligned} \tag{5}$$

and substituting (5) and (3) in (2),

$$0 = RAk\epsilon^{kt} + LAk^2\epsilon^{kt} + \frac{1}{C}A\epsilon^{kt}$$
$$= A\epsilon^{kt}\left(Rk + Lk^2 + \frac{1}{C}\right) \tag{6}$$

Neither A nor ϵ^{kt} can be zero, for this would mean by (3) that the current is zero for all values of time, and the problem has no significance. Therefore it follows from (6) that the remaining factor is zero, so that

$$Lk^2 + Rk + \frac{1}{C} = 0$$

or

$$k = \frac{-R \pm \sqrt{R^2 - \frac{4L}{C}}}{2L}$$

from which

$$\left.\begin{aligned} k_1 &= \frac{-R + \sqrt{R^2 - \frac{4L}{C}}}{2L} \\ k_2 &= \frac{-R - \sqrt{R^2 - \frac{4L}{C}}}{2L} \end{aligned}\right\} \tag{7}$$

Substituting (7) in (4),

$$i = A_1\epsilon^{\frac{-R+\sqrt{R^2-\frac{4L}{C}}}{2L}t} + A_2\epsilon^{\frac{-R-\sqrt{R^2-\frac{4L}{C}}}{2L}t} \tag{8}$$

In order to determine the two values of A indicated in (4), another relationship is needed. The voltage across the condenser is $E_c = q/C$ and from (1) is

$$E_c = E - iR - L\frac{di}{dt}$$
$$= E - (A_1\epsilon^{k_1t} + A_2\epsilon^{k_2t})R - L(A_1k_1\epsilon^{k_1t} + A_2k_2\epsilon^{k_2t}) \tag{9}$$

Since the current through an inductance cannot change instantaneously, it follows that when $t = 0$, $i = 0$. At the instant the switch S is closed the entire impressed voltage appears across L and $E_c = 0$ when $t = 0$. Therefore, from (4)

$$A_1 + A_2 = 0 \tag{10}$$

and (9) becomes

$$0 = E - L(A_1k_1 + A_2k_2)$$

or

$$E = A_1L(k_1 - k_2)$$

since $A_1 = -A_2$. Substituting the values of k_1 and k_2 from (7) in the above we get

$$\left.\begin{aligned} A_1 &= \frac{E}{\sqrt{R^2 - \frac{4L}{C}}} \\ A_2 &= -\frac{E}{\sqrt{R^2 - \frac{4L}{C}}} \end{aligned}\right\} \tag{11}$$

Inserting these values of A in (8), the equation of the current is

$$i = \frac{E}{\sqrt{R^2 - \frac{4L}{C}}}\epsilon^{-\frac{Rt}{2L}}\left\{\epsilon^{\frac{1}{2L}\sqrt{R^2-\frac{4L}{C}}\cdot t} - \epsilon^{-\frac{1}{2L}\sqrt{R^2-\frac{4L}{C}}\cdot t}\right\} \tag{12}$$

Equation (12) can be expressed in hyperbolic form from the relation $\sinh x = \frac{\epsilon^x - \epsilon^{-x}}{2}$, and an alternate expression for (12) will be

$$i = \frac{2E}{\sqrt{R^2 - \frac{4L}{C}}}\epsilon^{-\frac{Rt}{2L}} \sinh \frac{1}{2L}\sqrt{R^2 - \frac{4L}{C}}\cdot t \tag{13}$$

Making the above substitutions in (9) with the relation $\cosh x = \frac{\epsilon^x + \epsilon^{-x}}{2}$, the voltage across the condenser becomes

$$E_c = E - \frac{E\epsilon^{-\frac{Rt}{2L}}}{\sqrt{R^2 - \frac{4L}{C}}}\left\{R \sinh \frac{1}{2L}\sqrt{R^2 - \frac{4L}{C}}\cdot t + \right.$$
$$\left.\sqrt{R^2 - \frac{4L}{C}} \cosh \frac{1}{2L}\sqrt{R^2 - \frac{4L}{C}}\cdot t\right\} \tag{14}$$

Both (13) and (14) contain the quantity $\sqrt{R^2 - \frac{4L}{C}}$ and since

R, L, and C may have any value, three possibilities arise:

$$\text{Case I,}\quad R^2 > \frac{4L}{C}$$

$$\text{Case II,}\quad R^2 = \frac{4L}{C}$$

$$\text{Case III,}\quad R^2 < \frac{4L}{C}$$

Case I is taken care of by (13) and (14) and is illustrated in Fig. 67 for a typical case. The current through the circuit ultimately becomes zero and the voltage across the condenser becomes equal to the impressed voltage.

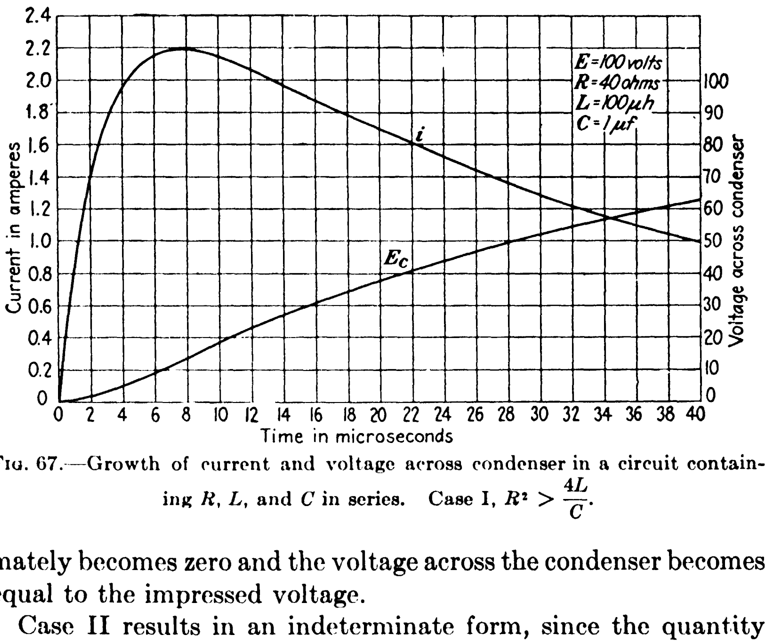


FIG. 67.—Growth of current and voltage across condenser in a circuit containing R, L, and C in series. Case I, $R^2 > \dfrac{4L}{C}$.

Case II results in an indeterminate form, since the quantity under the radical becomes zero when $R^2 = 4L/C$, so that (13) becomes zero divided by zero. Equation (12) may be written

$$i = \frac{E}{2L}\epsilon^{-\frac{Rt}{2L}}\left[\frac{\epsilon^{\frac{1}{2L}\sqrt{R^2-\frac{4L}{C}}\cdot t} - \epsilon^{-\frac{1}{2L}\sqrt{R^2-\frac{4L}{C}}\cdot t}}{\frac{1}{2L}\sqrt{R^2-\frac{4L}{C}}}\right]$$

The problem is to evaluate the quantity in brackets, which is of the form

$$\frac{\epsilon^{st} - \epsilon^{-st}}{s}$$

when $s = 0$. This can be done by differentiating the numerator and denominator with respect to s as follows:

$$\frac{dN/ds}{dD/ds} = \left[\frac{t\epsilon^{st} + t\epsilon^{-st}}{1}\right]_{s=0} = 2t$$

Substituting this value for the bracketed expression in the above equation for the current, gives us

$$i = \frac{Et}{L}\epsilon^{-\frac{Rt}{2L}} \tag{15}$$

which is the expression for the current in a circuit such as Fig. 66 when $R^2 = 4L/C$. This is often referred to as the *critical case,* in that the circuit resistance is just sufficient to prevent the current from oscillating. Values of resistance equal to, or greater than, this will cause the current to be aperiodic as shown in Fig. 67. The critical resistance is analogous to the case of a ballistic galvanometer, which, when shunted by a resistance of the proper value (also termed the critical resistance in this case), will just return to zero in the shortest possible time without overshooting the mark.

In a similar manner the voltage across the condenser when $R^2 = 4L/C$ is found to be

$$E_c = E\left[1 - \epsilon^{-\frac{Rt}{2L}}\left(\frac{Rt}{2L} + 1\right)\right] \tag{16}$$

Case III results in a negative quantity under the radical when $R^2 < 4L/C$, which can then be written

$$\sqrt{R^2 - \frac{4L}{C}} = j\sqrt{\frac{4L}{C} - R^2}$$

From the relations existing between the hyperbolic and circular functions, *viz.,*

$$\cosh jx = \cos x$$
$$\sinh jx = j \sin x$$

(13) becomes

$$i = \frac{2E}{j\sqrt{\frac{4L}{C} - R^2}} \epsilon^{-\frac{Rt}{2L}} \sinh j\frac{1}{2L}\sqrt{\frac{4L}{C} - R^2} \cdot t$$

$$= \frac{2E}{\sqrt{\frac{4L}{C} - R^2}} \epsilon^{-\frac{Rt}{2L}} \sin \frac{1}{2L}\sqrt{\frac{4L}{C} - R^2} \cdot t \qquad (17)$$

Likewise, the voltage across the condenser in this case will be

$$E_c = E - \frac{E\epsilon^{-\frac{Rt}{2L}}}{\sqrt{\frac{4L}{C} - R^2}} \left\{ R \sin \frac{1}{2L}\sqrt{\frac{4L}{C} - R^2} \cdot t + \sqrt{\frac{4L}{C} - R^2} \cos \frac{1}{2L}\sqrt{\frac{4L}{C} - R^2} \cdot t \right\} \qquad (18)$$

Equation (17) shows that the current is of the form of the product of a sine term and a term which decreases exponentially

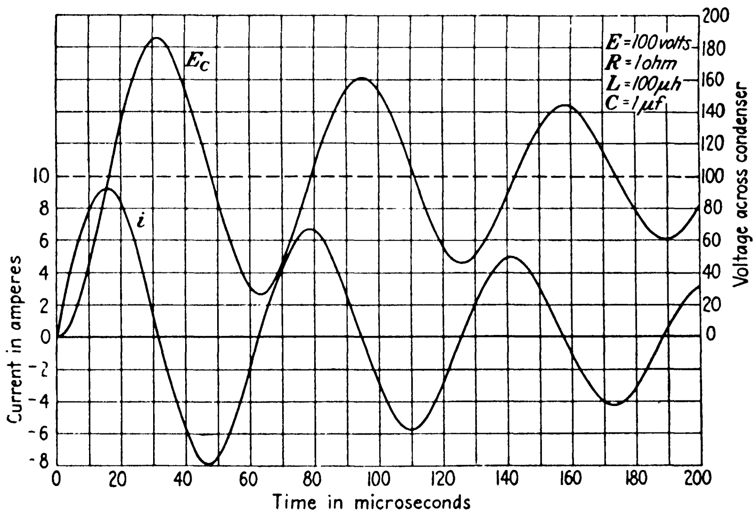


FIG. 68.—Current and condenser voltage in a circuit containing R, L, and C in series when a continuous voltage is impressed. Case III, $R^2 < \frac{4L}{C}$.

with time. It is plotted in Fig. 68, together with the voltage across the condenser. It will be noted that the latter rises to a value of almost twice the impressed voltage if the circuit resistance is small. With R zero, the maximum value of E_c is $2E$.

This should be taken into consideration in condenser tests where the usual test procedure is to impress a high value of direct-current voltage on the condenser. If the internal resistance of the source is small, the inductance of the source, including the leads, may constitute an oscillatory circuit and subject the condenser to a peak voltage which will be nearly twice the supposed test voltage.

The sine function of time which appears in (17) is equivalent to sin ωt or

$$\omega = \frac{1}{2L}\sqrt{\frac{4L}{C} - R^2}$$

and the frequency of oscillation will be

$$f = \frac{1}{2\pi}\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}} \tag{19}$$

It will be observed that the frequency of oscillation is affected by the resistance as was the resonant frequency in the parallel resonance case. However, the expressions are not identical as will be seen by comparing the expression with (16) of Chap. II.

If R is small compared to $4L/C$, the approximate expression for the current may be written

$$i = \frac{E}{\sqrt{L/C}}\epsilon^{-\frac{Rt}{2L}} \sin \frac{t}{\sqrt{LC}} \tag{20}$$

The maximum value of current, neglecting the damping factor $\epsilon^{-\frac{Rt}{2L}}$ during the first quarter of the cycle, is given by

$$I_{max} = \frac{E}{\sqrt{L/C}} \tag{21}$$

By the definition of impedance, it is seen that the term $\sqrt{L/C}$ is of the nature of an impedance. It is called the *surge impedance*. This factor limits the magnitude of the surge current that would flow when a voltage is suddenly applied to a circuit composed of L and C in series, such as a transmission line on open circuit.

40. Discharge of a Condenser through an Inductance and Resistance.—If a condenser is initially charged and then allowed to discharge through a coil having inductance and resistance by closing the switch S in Fig. 69, the current at any instant will be given by

$$0 = iR + L\frac{di}{dt} + \frac{1}{C}\int idt$$

This equation, upon differentiating, is identical with (2), and hence its solution will be the same and will not be repeated here, the only difference being in the boundary conditions. The important case is where $R^2 < 4L/C$ and the expression for the current under this condition is

$$i = -\frac{2E_c}{\sqrt{\frac{4L}{C} - R^2}}\epsilon^{-\frac{Rt}{2L}} \sin \frac{1}{2L}\sqrt{\frac{4L}{C} - R^2} \cdot t \qquad (22)$$

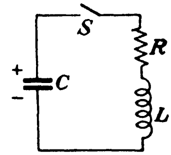


FIG. 69.—Condenser C will discharge through L and R when switch S is closed.

where E_c is the initial voltage of the condenser. It will be observed that this expression is identical in form with (17). The negative sign merely means that the current is opposite in direction to the charging case.

Formerly, practically all radio transmitting stations employed the oscillatory discharge of a condenser through an inductance as the means of producing high-frequency currents. The antenna was usually coupled inductively to the circuit. Instead of the switch S, some form of spark gap was used. The condenser was charged by means of either a high-voltage step-up transformer or an induction coil. The antenna current consisted of a series of damped oscillations, one wave train being produced by each spark across the gap. These high-frequency currents set up an alternating electromagnetic and an alternating electrostatic field surrounding the antenna which constitute electromagnetic waves. These waves are radiated into space with the velocity of light, as will be discussed later.

The coupling of the antenna circuit to the primary oscillating circuit causes a complex oscillation in both circuits, which will be considered in the following section.

41. Free Oscillations in Coupled Circuits.—If the condenser C_1 is assumed to be charged in the coupled circuit of Fig. 70 and is then allowed to discharge, the primary and secondary currents at any instant will be given by

$$0 = i_1R_1 + L_1\frac{di_1}{dt} + M\frac{di_2}{dt} + \frac{1}{C_1}\int i_1dt \qquad (23)$$

$$0 = i_2R_2 + L_2\frac{di_2}{dt} + M\frac{di_1}{dt} + \frac{1}{C_2}\int i_2dt \qquad (24)$$

Differentiating (23) and (24), and letting D stand for d/dt and D^2 for d^2/dt^2, etc.,

$$0 = MD^2i_2 + \left(L_1D^2 + R_1D + \frac{1}{C_1}\right)i_1 \qquad (25)$$

$$0 = MD^2i_1 + \left(L_2D^2 + R_2D + \frac{1}{C_2}\right)i_2 \qquad (26)$$

To eliminate i_2, apply the operator $\left(L_2D^2 + R_2D + \frac{1}{C_2}\right)$ to (25) and MD^2 to (26). These become,

$$0 = MD^2\left(L_2D^2 + R_2D + \frac{1}{C_2}\right)i_2 + \Big[L_1L_2D^4 + (R_1L_2 + R_2L_1)D^3 + \left(\frac{L_1}{C_2} + R_1R_2 + \frac{L_2}{C_1}\right)D^2 + \left(\frac{R_2}{C_1} + \frac{R_1}{C_2}\right)D + \frac{1}{C_1C_2}\Big]i_1 \qquad (27)$$

$$0 = MD^2\left(L_2D^2 + R_2D + \frac{1}{C_2}\right)i_2 + M^2D^4i_1 \qquad (28)$$

Subtracting (28) from (27) and multiplying by C_1C_2 gives

$$C_1C_2(L_1L_2 - M^2)D^4i_1 + C_1C_2(R_1L_2 + R_2L_1)D^3i_1 + (C_1L_1 + C_1C_2R_1R_2 + C_2L_2)D^2i_1 + (C_1R_1 + C_2R_2)Di_1 + i_1 = 0 \qquad (29)$$

By a similar procedure an identical equation can be obtained for i_2. The above expression is a linear differential equation of the fourth order and its complete solution involves the solution of a fourth-degree algebraic equation. As we are more interested in the behavior of the circuit from the standpoint of the frequencies of oscillation, an exact solution for the current will not be attempted.

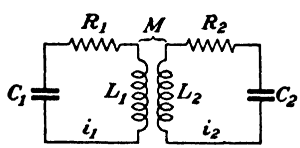


FIG. 70.—Two oscillatory circuits magnetically coupled.

If the circuit resistance is small, the frequency of oscillation is but little affected, as previously shown. Neglecting the resistances and expressing M in terms of the coefficient of coupling k, (29) becomes

$$C_1C_2L_1L_2(1 - k^2)\frac{d^4i}{dt^4} + (C_1L_1 + C_2L_2)\frac{d^2i}{dt^2} + i = 0 \qquad (30)$$

Assuming the resistances to be negligible enables the current to be expressed by $i = I \sin \omega t$. The derivatives of the current are

$$\frac{d^2i}{dt^2} = -I\omega^2 \sin \omega t$$

$$\frac{d^4i}{dt^4} = I\omega^4 \sin \omega t$$

Substituting these values in (30), we get

$$C_1C_2L_1L_2(1 - k^2)I\omega^4 \sin \omega t - (C_1L_1 + C_2L_2)I\omega^2 \sin \omega t + I \sin \omega t = 0 \qquad (31)$$

or

$$\frac{1 - k^2}{f_1^2 f_2^2}f^4 - \left(\frac{1}{f_1^2} + \frac{1}{f_2^2}\right)f^2 + 1 = 0 \qquad (32)$$

where f_1 and f_2 are the resonant frequencies of the primary and secondary.

Solving (32) for f,

$$f = \sqrt{\frac{f_1^2 + f_2^2 \pm \sqrt{(f_1^2 - f_2^2)^2 + 4k^2 f_1^2 f_2^2}}{2(1 - k^2)}} \qquad (33)$$

which is identical with (29) of Chap. IV, since resistance has been assumed to be negligible in both cases.

From (33) it is seen that there will be two frequencies present, the higher value pertaining to the plus sign under the radical and the lower frequency being given by the minus sign. If the natural frequencies of the primary and secondary are equal ($C_1L_1 = C_2L_2$) so that $f_1 = f_2 = f_0$, (33) becomes

$$f = \frac{f_0}{\sqrt{1 \pm k}} \qquad (34)$$

As the coefficient of coupling approaches zero, the two frequencies of oscillation approach each other. The two frequencies produce beats in the resultant oscillations as illustrated in Fig. 71 which depicts the currents in the primary and secondary circuits when they both are tuned to the same frequency. The larger the coefficient of coupling, the greater the difference in the two frequencies, and the shorter the interval between beats. An

excellent illustration of the behavior of a freely oscillating coupled circuit is given by the coupled pendulum shown in Fig. 72. The two pendulums, individually adjustable as to length, are sup-

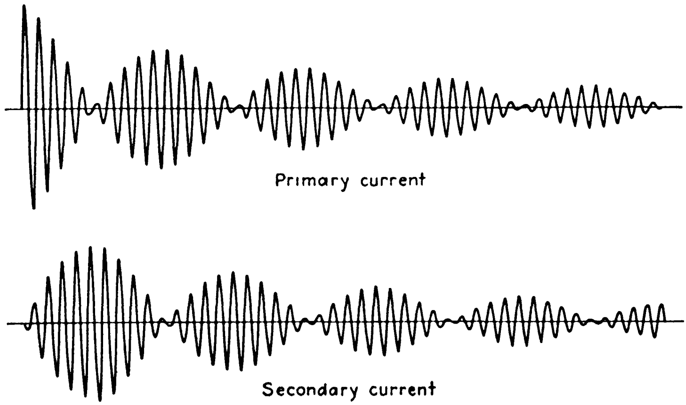


FIG. 71.—Currents in the primary and secondary of a coupled circuit when both circuits are tuned to the same frequency.

ported by the horizontal string, the tension in which corresponds to the degree of coupling between them. The pendulum initially disturbed functions as the primary and communicates its motion to the secondary, their motions being analogous to the current. Beats are produced in their oscillations in exactly the same manner as in the electrical case. By changing the lengths of the pendulums the effects of detuning one circuit or the other may be observed.

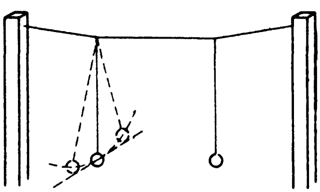


FIG. 72.—Coupled pendulum model illustrating behavior of coupled circuits.

42. Spark Transmitters.—This term is applied to transmitters which produce radio-frequency currents by means of the oscillatory discharge of a condenser through an inductance in series with a spark gap. Transmitters of this type are practically obsolete and only a brief discussion of them will be given. The principles used are still employed in high-frequency demonstrations using Tesla coils and similar apparatus. High-frequency currents produced in this fashion are occasionally used in induction furnaces and in vacuum-tube manufacture when it is necessary to heat the elements inside of the glass bulb by eddy currents during the evacuating process.

This method of producing high-frequency current is usually a fruitful source of radio interference so that other methods are usually employed.

The typical circuit of a spark transmitter is shown in Fig. 73. When the key is closed the low-frequency alternating voltage E is stepped up to about 10,000 or 15,000 volts by means of a suitable transformer and impressed on the condenser C_1. When the secondary voltage of the transformer rises to the point where it is sufficient to break down the gap, the condenser discharges through L_1 and induces a voltage in L_2, which has been adjusted so that the antenna circuit L_2C_2 is in resonance with the primary oscillating circuit. The antenna current will have the general appearance of the secondary current in Fig. 71, which means that two frequencies are being radiated. This is highly objectionable as it causes the station to occupy too wide a frequency band which would produce considerable interference with other stations, and also because the total energy is subdivided, a portion being radiated at one frequency and the remainder on the other, with consequent reduction in the range of transmission. The frequency band may be narrowed in width by reducing the coupling between L_1 and L_2, but weak coupling must be used before very much improvement is produced, which appreciably reduces the radiated power. If the primary circuit could be prevented from oscillating after all of its energy has been transferred to the secondary, the beats present in the secondary current would be eliminated and it would then oscillate at a single frequency which would be damped out at a rate governed by the antenna constants. This can be accomplished by the use of a gap whose resistance rises to a very high value as the amplitude of the primary current diminishes and which remains sufficiently high so that it is not re-ignited by the voltage induced from the secondary coil L_2. A rotary gap consisting of a toothed wheel rotating between fixed electrodes is of assistance in accomplishing this result. In some commercial types of transmitters the gap wheel was mounted directly on the shaft of the alternator which supplied power for the set. There would be as many teeth on the wheel as

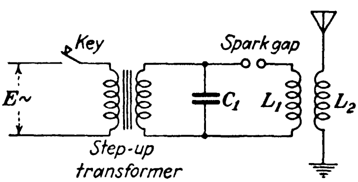


FIG. 73.—Circuit diagram of a spark transmitter.

there are poles on the alternator so that synchronous operation would result, producing one spark per half cycle. Another type, known as the "quenched gap" is more satisfactory. It consists of a number of flat copper disks, separated by insulating gaskets forming a series of airtight chambers in which the sparks take place. The distance between each gap is short (from 0.01 to 0.02 in.) with a sparking area of several square inches. This construction results in a rapid deionizing so that the gap resistance rises rapidly as the current through it diminishes, which causes a rapid damping of the primary, allowing the secondary to oscillate without being reacted upon by the primary circuit. As a further aid to rapid damping in the primary, the ratio of C_1 to L_1 should be kept as large as possible. The damping factor $\epsilon^{-\frac{Rt}{2L}}$ will become greater with a fixed value of R as L is reduced. The ideal condition would be for the primary to be discharged in a single large aperiodic impulse. All of its energy would then be transferred to the secondary during this single impulse, thereafter allowing the secondary to oscillate freely without interreaction between the two circuits. Under this condition there is no need of tuning the primary circuit. Transmitters approaching this mode of operation were known as "impulse transmitters." An aperiodic primary could be obtained by increasing the primary resistance so that R^2 is greater than $4L/C$, but this would cause most of the energy to be dissipated in I^2R loss. Even with a gap capable of good quenching action the coupling between the primary and secondary oscillating circuits must not be too tight, or re-ignition of the gap will take place.

The step-up transformer used to charge the condenser C_1 must be designed with poor regulation as the breakdown of the gap is virtually a short circuit of the transformer secondary. In addition to the excessive current that would flow, there would be a tendency to maintain an arc across the gap which would prevent the gap from acting as a switch whose opening and closing at the proper points of the cycle allow the condenser to charge and discharge. This poor regulation on the part of the transformer is usually obtained by designing it to have high leakage flux between the primary and secondary. This can be accomplished in a core-type construction by winding the primary on one leg and the secondary on the other, sometimes augmenting the leakage by a magnetic shunt across the primary. A suitable external

reactance in series with the primary can accomplish the same result. A considerable portion of effective series reactance is obtained in the relatively high internal synchronous reactance present in the 500-cycle alternators usually employed with commercial types of spark transmitters.

This high value of frequency was used so as to make the received signals clearly audible through the noise due to atmospheric disturbances, such as static, and to take advantage of the greater sensitivity of the human ear at the higher audio fre-

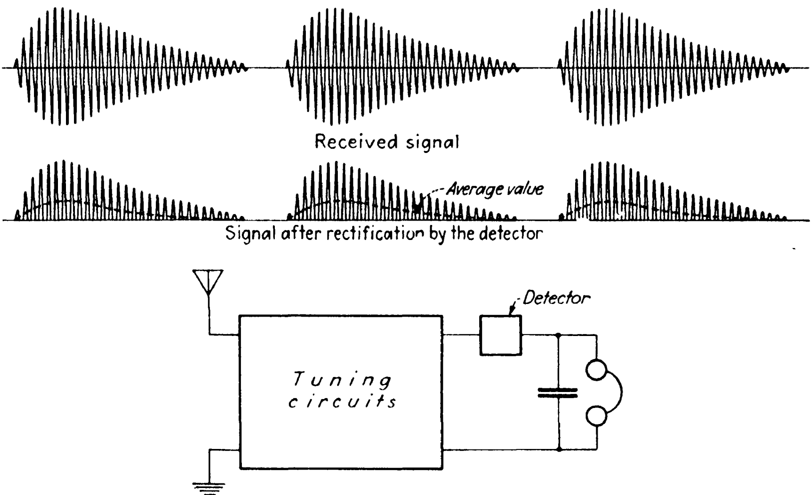


FIG. 74.—Received signal before and after rectification by the detector.

quencies. With either the synchronous rotary gap or the quenched gap, one spark per half cycle is produced giving rise to one oscillation or wave train in the antenna for each spark. A similar oscillatory voltage is induced in the receiving antenna by the impinging electromagnetic wave. This oscillation passes through the tuned circuits of the receiving set and is impressed on the detector. The ideal detector would be a perfect rectifier which would remove the lower half of the received oscillations as shown in Fig. 74. If a pair of telephone receivers are connected in series with the rectifying device, the average value of current through them would be similar to the dotted curve. The by-pass condenser across the telephone receivers serves to integrate the succession of rectified half waves of the train into

the single dotted impulse. Each wave train produces a single impulse in the receivers so that a 500-cycle spark transmitter producing two wave trains per cycle will cause a 1000-cycle tone to be heard in the telephone receivers whenever the key at the transmitter is depressed. The dots and dashes of the telegraphic code can thus be formed.

Where only 60 cycles was available, a nonsynchronous rotary gap was used. This produced a number of sparks per cycle, enabling a desirable high-pitched note to be obtained. The number of sparks per second depended upon the speed of the disk and the number of teeth. The setting of the gap electrodes in this case had to be short enough to permit the gap to break down at voltages much lower than the peak value, otherwise a ragged note would be produced. This type of gap was popular with the amateur radio fraternity prior to the introduction of vacuum-tube transmitters using continuous waves.

The power requirements of a spark transmitter can be determined from the energy stored in a condenser which is given by

$$W = \tfrac{1}{2}CE^2 \text{ joules} \tag{35}$$

The power in watts required by the condenser, which is equal to the necessary rating of the transformer, will be

$$p = \frac{nCE^2}{2} \tag{36}$$

where C is the capacitance of the transmitting condenser in farads, n is the number of charges of the condenser per second, and E is the voltage across the condenser at the instant the gap breaks down. If one spark occurs at each alternation of the power supply, n will be equal to $2f$.

43. Radio Interference from Electrical Sparks.—From what has been said in the preceding sections on oscillating circuits it will be seen that whenever an electrical circuit is made or broken, an electric oscillation may be set up. All circuits possess inductance and capacitance—either distributed, or with respect to ground in the case of the latter—so that these circuits are all potential transmitters. Fortunately, the transmission range of these disturbances is usually not more than a few hundred feet, but in some cases the oscillations may be transmitted along power circuits for considerable distances, and radiate energy

which may be picked up by a sensitive receiving set anywhere in the immediate vicinity of the conductors. The ignition system of every automobile is a mobile transmitter with a transmitting range of several hundred yards at frequencies of 10^7 cycles and higher, unless it is equipped with its own receiving set, in which case filters and suppressors must be employed in the electrical system to reduce the noise produced by these unwanted oscillations. The suppressors used in the spark-plug leads are high resistances which not only render the circuit nonoscillatory but also reduce the steepness of the impulse. A single aperiodic impulse is capable of setting up an oscillation in the resonant circuits of a receiving set if the wave front of the impulse is steep enough. In other words, merely preventing oscillations from occurring in electrical circuits is no assurance of freedom from noise in adjacent receiving sets. From the standpoint of radio interference direct-current circuits are usually much noisier than alternating-current circuits, largely because of the continuous interruption of portions of armature circuits in the process of commutation.

44. Limitations of Spark Transmitters.—In addition to the possibility of radiating energy at two adjacent frequencies due to improper adjustment, spark transmitters produce a great deal more interference with each other than do the continuous-wave transmitters which have replaced them. This is due to the damping present in their transmitted waves; the more rapidly the oscillations in the antenna circuit die out, the more difficult it becomes to tune out the particular transmitter. The reason for this can be best understood by the exactly analogous case in sound. Suppose we have a stringed instrument, such as a piano, in a room and produce a sustained musical tone by some other instrument. The string on the piano which is tuned to the same frequency as this tone will respond, while those closely adjacent will respond to a much lesser degree. The sound waves produced by the sustained tone are impinging on all the other strings as well, but get into interference with them and oppose their natural vibratory motion as often as they assist it, so that the motion of the other strings is negligibly small. The initial impulse of the sound wave struck all of the strings, which would have set them into vibration at their own natural frequencies if it were not for the interference caused by the subsequent timed impulses of

the disturbing musical tone. However, if a loud noise is produced in the room, such as a shout or the dropping of a heavy object, all of the piano strings are set into vibration. This is exactly similar to the effect of static on a number of receiving sets tuned to different frequencies. The irregular impulses of static set the various receivers into oscillation by impact excitation and each receiver oscillates at a frequency dictated by its own circuit constants. Consequently, static and other similar irregular electrical impulses will be received regardless of the tuning adjustments of the set. Very strongly damped oscillations from a spark transmitter will act in the same fashion and will be heard in the receiver even when it is tuned to a frequency remote from the interfering oscillations. Use was made of this in transmitting distress signals. The radio operator would increase the coupling in the transmitter which would broaden the transmitted wave considerably, since it would now contain two peaks and the damping in the antenna circuit would be increased, owing to the added resistance reflected from the primary oscillating circuit. The possibility of the signals being heard was thereby increased. Much of the electrical interference discussed in the preceding section is usually strongly damped in character so that the noise is spread over a considerable portion of the tuning dial without a very well defined maximum.

Another point of view which is of use in understanding the inability of selective tuned circuits to discriminate against irregular impulses is to consider the latter to be resolved into a Fourier series of an infinite number of harmonics. Tuning the receiving circuit to various frequencies merely resonates it to one of these harmonics. The absence of static at the very high frequencies (5×10^7 cycles and higher) is then explained by the almost negligible magnitudes of the constituent harmonics in this range of frequencies.

As the damping in spark signals becomes less, the selectivity at the receiving end improves. With zero damping we have a "continuous wave" and the selectivity then depends only upon the sharpness of resonance of the tuned circuits employed in the receiver. Zero damping is impossible with spark excitation owing to the resistance present in the antenna circuit.

The energy radiated by the electromagnetic waves can be accounted for by considering the antenna resistance to be

increased by an amount sufficient to take care of the radiated energy. This radiation resistance of the antenna augments the ohmic resistance and the resistance due to dielectric losses in the electrostatic field of the antenna. The total antenna resistance will be

$$R_a = R_o + R_d + R_r \tag{37}$$

the three components being ohmic, dielectric, and radiation resistance, respectively.

The amount of damping present in the antenna circuit may be determined by the sharpness of resonance of a suitable tuned-circuit indicator, such as a wave meter.[1] The latter is merely a coil shunted by a variable condenser, with a suitable indicating instrument A in series, as shown in Fig. 75. A hot-wire or a thermocouple instrument is usually used.

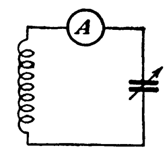


FIG. 75.—Wave meter circuit.

The coil of the wave meter is loosely coupled to the circuit to be measured and resonance is indicated by adjusting the condenser for maximum deflection of the indicating instrument. The device may be calibrated to read the frequency, or else the wave length in meters. The relation between these two is given by

$$\lambda = \frac{c}{f} = 1885\sqrt{LC} \tag{38}$$

where λ is the wave length in meters, c the velocity of light in meters per second (approximately 3×10^8), f is in cycles, and L and C are microhenrys and microfarads, respectively.

If the circuit is strongly damped, the setting of the wave meter condenser for maximum current will be quite broad. In fact, when the wave meter is coupled to a primary oscillating circuit having good quenching action, it is almost impossible to determine the frequency with any accuracy.

Spark transmitters are seriously limited as to power output at radio frequencies much above 1500 kc. As the operating frequency rises and C_1 stays fixed, L_1 will have to be reduced until it is too small to furnish adequate coupling with L_2. Reducing C_1 will reduce the power in direct proportion unless the transformer voltage is raised at the same time, which introduces insulation difficulties and complicates the spark-gap design.

[1] *Bur. Standards Circ.* 74, p. 187.

Problems

1. A 4-μf condenser is connected across a 1000-volt source of direct current. The total inductance of the leads and source is 2 μh and the resistance of the circuit is negligible. What will be the maximum instantaneous current flowing through the condenser? What will be the maximum instantaneous voltage across the condenser? What is the minimum value of resistance to be placed in series with the condenser so as to prevent the condenser voltage from exceeding 1000 volts?

2. In the aperiodic case of the growth of current in a circuit containing R, L, and C in series, prove that the current reaches its maximum value in a time

$$t = \frac{L}{\sqrt{R^2 - \frac{4L}{C}}} \log_\epsilon \frac{R + \sqrt{R^2 - \frac{4L}{C}}}{R - \sqrt{R^2 - \frac{4L}{C}}}$$

3. A surge generator used to produce artificial lightning for test purposes is constructed of 20 condensers each having a capacitance of 0.25 μf. The circuit is so arranged that the condensers are all charged in parallel to 100,000 volts and then are connected in series and discharged through a circuit composed of a resistance of 260 ohms in series with an inductance of 200 μh. How long a time will be required for the discharge current to reach its maximum value?

4. What will be the maximum value of the current in Problem 3?

5. The primary oscillating circuit of a 500-cycle spark transmitter is adjusted to oscillate at a frequency of 500 kc. The capacitance of the condenser is 0.01 μf and the effective resistance of the circuit is 1.5 ohms, which may be assumed constant. If the condenser is charged to 15,000 volts when the gap breaks down, what will be the effective value of the current in the circuit? What power must the transformer supply? What is the maximum value of the current, neglecting resistance?

CHAPTER VI

FUNDAMENTAL PROPERTIES OF VACUUM TUBES

45. Introduction.—Much of the rapid progress in electrical communication in recent years has been due to the thermionic vacuum tube. In addition to their applications in this field they have become an indispensable tool in many types of research work. Thermionic tubes are also being used to an increasing extent in various industrial control applications, particularly in conjunction with photocells.

The physics of the vacuum tube has been rather extensively treated in the literature and only a brief discussion will be given here. For a more detailed treatment containing an extensive bibliography the reader is referred to Chaffee's "Theory of Thermionic Vacuum Tubes."

The ordinary type of vacuum tube consists of a cathode capable of emitting electrons when heated, an anode or plate which is held at a positive potential so as to attract the electrons, and some means of controlling their flow, usually by one or more electrodes placed between the anode and cathode. These electrodes are enclosed in a suitable evacuated container. Tubes are classified as diodes, triodes, tetrodes, pentodes, etc., depending upon the number of electrodes present. Most of these types are very highly evacuated, although some contain small amounts of gas or metallic vapor so as to produce certain operating characteristics. Thyratrons and grid-glow tubes are examples of the latter.

46. History of the Thermionic Tube.—For several years, beginning in 1882, Elster and Geitel had studied the conductivity of a gas in the vicinity of heated solids and flames, and while in some of their later work they actually used a two-element vacuum tube consisting of a plate and an electrically heated carbon filament, they did not mention the possibilities of the device as an alternating-current rectifier, although they had noted its unilateral conductivity. The use of alternating current was quite limited at this early date so that the need or uses of rectifiers were not appreciated.

Edison, in 1883, observed that if a plate sealed in the incandescent lamp, which he had recently developed, were connected to the positive end of the filament through a galvanometer, a current would flow. If the galvanometer lead was connected to the negative end of the filament, no current flowed. This phenomenon has been known as the *Edison effect*. In 1899, J. J. Thomson showed that this current was due to the travel of electrons from the filament to the plate. J. A. Fleming[1] in 1904 patented the two-electrode vacuum tube, or "Fleming valve," as it was called, as a detector of high-frequency oscillations through the rectifying action of the device.

In 1907[2] Lee de Forest made an important improvement in the vacuum tube in the form of a third electrode or "grid" introduced between the filament and the plate. This provided a means of fundamental importance for controlling the flow of electrons in the "audion," as he called the device. This control electrode enables the three-electrode tube, or triode, to perform its functions as an amplifier and oscillator.

Further improvements in industrial research laboratories under Arnold and Langmuir followed, notably the use of a high vacuum and the development of better types of filament. W. Schottky,[3] in 1919, suggested the use of a second grid as an electrostatic screen between the plate and control grid. This "screen-grid" tube was further developed by Hull and Williams[4] and began to be used in broadcast receiving sets in 1928. Prior to this time the requirements of radio reception had been met by a comparatively few types of general-purpose tubes whose chief differences lay in the matter of filament voltage and power. Shortly after the advent of the screen-grid tube other special-purpose, multielectrode tubes began to appear in increasing numbers, and for a variety of filament voltages, so that hundreds of types are now on the market. A number of these merely combine two or more groups of elements within a single bulb—such as two triodes, or two diodes and a triode—which enables a single tube to perform the functions of two or more individual tubes. In 1935, a new construction was introduced which

[1] British patent 24,850, Nov. 16, 1904.
[2] U. S. patent 879,532, filed Jan. 29, 1907.
[3] *Arch. Elektrot.*, vol. 8, p. 299, 1919.
[4] *Phys. Rev.*, vol. 27, p. 433, 1926.

substituted an evacuated metal container for the glass bulb. This will ultimately bring about a reduction in the number of tube types as many of the older varieties will not be duplicated in the "all-metal" design.

The first patent disclosure of the oscillating properties of the vacuum tube by Meissner[5] stimulated activity in the development of tubes capable of handling considerable amounts of power. Credit for this discovery is also attributed to Armstrong and de Forest in America, and Franklin and Round in England. The art was developing so rapidly that it is difficult to determine where the credit belongs, as a considerable period of experimental work usually precedes the date of patent application.

The engineers of the American Telephone and Telegraph Company and Western Electric Company used about three hundred 25-watt tubes in an experimental test with radio telephony at the Naval Station at Arlington, Va., in 1915. They succeeded in being heard in Paris and also in Honolulu. Encouraged by these results the research divisions of these organizations began experiments with a view of increasing the power rating of the tubes, as the use of a large number of small tubes in parallel as oscillators, modulators, and amplifiers is inefficient and rather inconvenient. In the ordinary type of construction using an internal plate, all of the heat developed at the plate by electronic bombardment must be dissipated by radiation, which limits the maximum rating of the tube to about 2 kw. The development of a method of sealing copper to glass by W. G. Housekeeper enabled a new form of construction to be used. The plate was a copper tube welded into a glass base and formed a portion of the containing envelope. This design enables the plate to be water-cooled and permits the construction of tubes having a rating of 500 kw. A water-cooled grid of tubular construction is also used in this size.

47. Emission of Electrons.—Electrons may be regarded as minute negatively charged particles that are one of the constituents of all matter. They have a mass of 9.035×10^{-28} gram and a charge of 1.59×10^{-19} coulomb. The giving off of electrons from solid bodies may be accomplished by raising the temperature so that the free electrons of the material are ejected

[5] German patent, 291,604, Apr. 10, 1913.

and escape from the surface tension of the body. This process is called *thermionic emission.* Only those electrons possessed of a velocity greater than a certain minimum value are able to overcome this surface restraint and succeed in escaping. The amount of work w represented by this escape is known as the *thermionic work function* and represents the minimum amount of kinetic energy the electron must have in order to pass through the surface. The work w is ordinarily expressed in terms of a potential V such that V times the charge e on the electron is equal to w, or

$$Ve = w = \tfrac{1}{2}mv^2 \tag{1}$$

where m is the mass of the electron and v is the minimum velocity necessary to escape the surface restraint. The equivalent voltage V is called the *electron affinity* and is a measure of the difficulty that electrons experience in escaping from different materials. For tungsten it has a value of 4.52 volts.

The escaped electrons charge the space outside negatively and leave the emitter positively charged, so that an electrostatic field is produced which tends to force the electrons to return. They would all eventually return to the emitting body in the absence of any other positive charge in the vicinity. In the case of the ordinary type of vacuum tube the positive potential on the plate attracts the electrons, giving rise to a *thermionic current,* the magnitude of which depends upon the number of electrons per second reaching the plate. This migration of electrons from cathode to plate is equivalent to saying that a current is flowing *from plate to cathode* within the tube, since our conventional direction of current flow (which is a purely arbitrary assumption) happens to be opposite in direction to the actual migration of electrons through the circuit.

The higher the temperature of the cathode or filament, the greater the number of electrons which succeed in escaping. If the plate is sufficiently positive with respect to the filament, all of the emitted electrons will reach the plate and a further increase in plate potential will produce no increase in plate current. This saturation value of thermionic current per unit area of filament was shown by Richardson[6] to be

[6] O. W. RICHARDSON, "Emission of Electricity from Hot Bodies," Longmans, Green & Company, 1916.

$$I_s = AT^2\epsilon^{-\frac{b_0}{T}} \tag{2}$$

where T is the absolute temperature in degrees Kelvin, and b_0 and A are constants depending on the emitting material. In the case of pure metals A has the value of 60.2. The electron affinity of the material governs the value of b_0, which is 52,400 for pure tungsten.

If the plate potential is lowered, the intensity of the electrostatic field at the filament is reduced. The field at this point is the resultant of two opposite forces, namely, the potential of the plate, and the negative charges due to the swarm of electrons surrounding the filament. This cloud of electrons constitutes a *space charge* which causes the thermionic current to fall below

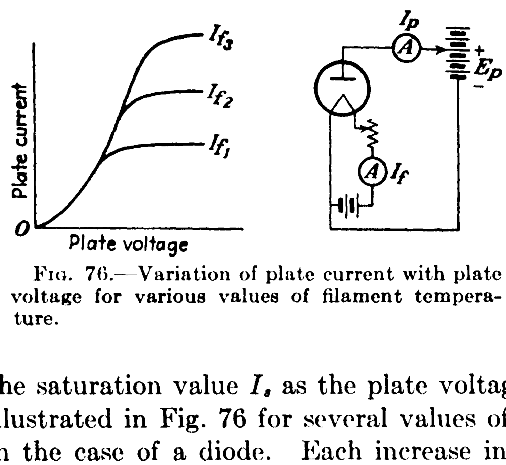


FIG. 76.—Variation of plate current with plate voltage for various values of filament temperature.

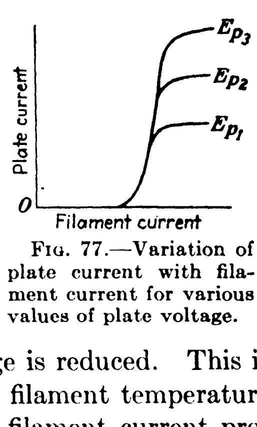


FIG. 77.—Variation of plate current with filament current for various values of plate voltage.

the saturation value I_s as the plate voltage is reduced. This is illustrated in Fig. 76 for several values of filament temperature in the case of a diode. Each increase in filament current produces a higher value of saturation current since there are more electrons available.

The effect of space charge is illustrated in Fig. 77, using the same circuit as before but making the filament current the independent variable while the plate voltage E_{p1} is held constant. Saturation again occurs as the filament temperature is increased, but it is now due to space charge as there are abundant electrons available. Raising the plate voltage to a value E_{p2} immediately increases the plate current to a higher saturation value. At values of plate current less than saturation where the current is limited by space charge, Child[7] has shown for the case where the

[7] *Phys. Rev.*, vol. 32, p. 498, 1911.

cathode and plate are infinite plane surfaces that the theoretical value of current is

$$I_p = 2.336 \times 10^{-6}\frac{E_p^{3/2}}{d^2} \text{ amp. per square centimeter of cathode} \quad (3)$$

where d is the distance between electrodes in centimeters. This expression shows that the plate current varies as the three-halves power of the plate voltage. The derivation assumes that the cathode is an equipotential surface and that the electrons emerge from it with zero velocity. For a cylindrical plate with the filament along the axis, Langmuir has shown that

$$I_p = 14.68 \times 10^{-6}\frac{E_p^{3/2}}{r} \text{ amp. per centimeter length of cathode} \quad (4)$$

where r is the radius of the plate in centimeters. The diameter of the filament is assumed to be negligible in comparison with the diameter of the plate. The general expression for the plate current where insufficient emission is not a factor is of the form

$$I_p = KE_p^{3/2} \quad (5)$$

where K is a constant determined by the geometry of the tube. At low values of plate voltage where the plate current is small, the velocity of emission and *ir* drop (if present) in the cathode will cause a deviation from the three-halves power and the characteristic will frequently follow a square law in this region. The potential of various portions of the filament with respect to the plate will vary along the length of the filament if *ir* drop is present in the latter. A similar deviation occurs at high values of plate current due to inadequate emission from the cathode as saturation is approached.

In addition to thermionic emission, electrons may be given off by a body if it is bombarded by electrons or ions moving with sufficient velocity. This is termed *secondary emission* and is of importance in a type of tube known as the dynatron. Usually secondary emission is a source of annoyance and has to be guarded against in tube design.

A third form of electronic emission occurs when electromagnetic radiation of sufficiently short wave length, such as x-rays and ultraviolet light, falls on a body. This is known as *photoelectric emission* and is the principle involved in photocells. The

alkaline metals potassium and caesium exhibit this phenomenon for wave lengths within the visible spectrum of light. The emitting surface is usually a combination of various materials chosen with the idea of improving the sensitivity within the visible spectrum. While photocells are a necessary adjunct to television and talking pictures, a detailed discussion of their theory and characteristics is beyond the scope of this chapter. For such information the reader is referred to "Photoelectric Cells and Their Applications" by Zworykin and Wilson, and to "Photoelectric Phenomena" by Hughes and DuBridge.

48. Types of Cathodes.—The cathode of a vacuum tube may be either in the form of a filament heated by the passage of current through it, as previously mentioned, or of the indirectly heated type. The construction of the latter usually consists of an elongated metal cylinder of small diameter inclosing a heating element. The heating element is a hairpin loop of tungsten wire enclosed in a suitable refractory material, or to secure quicker heating it is often coiled inside the cylinder and supported at each end by insulating plugs. This type of cathode is widely used in tubes for radio reception where only alternating current is available for heating purposes as it greatly reduces the objectionable hum that would be present when the filament type of tube is heated by alternating current.

The indirectly heated or equipotential type of cathode would be impractical if it were not for the existence of emitters capable of satisfactory emission at relatively low temperatures. It would be almost impossible to secure the high temperature required by tungsten or tantalum for reasonable emission by heating them indirectly unless a material were discovered for a heater which had a considerably higher melting point than either of these two.

In the larger types of thermionic devices, particularly in mercury-vapor rectifiers and thyratrons where a very large current is carried, the watts required to heat the filament would be an appreciable tax upon the overall efficiency of the device. If the radiation loss at the cathode can be reduced, the emission efficiency, expressed in amperes emission per watt of filament heating power, can be considerably increased. One way in which this can be accomplished is to use a ribbon type of filament coiled in a flat spiral in a form similar to a clock spring. In this way

each turn shields the other so that the heat loss by radiation is mostly from the outside surface of the outer turn. The radiation from the edges is small. This idea has been further developed in the case of equipotential cathodes by using a series of concentric polished nickel cylinders which surround the cathode. With this construction it is possible to bring the cathode to the proper operating temperature with only 2 or 3 per cent of the heating power that would be necessary if this heat shielding were not employed. From 0.5 to 1.5 amp. of emission per watt of filament power can be obtained. Small holes are provided in this "oven" through which the electrons escape.

This type of design is of no use in high-vacuum tubes owing to the effects of space charge. The electrostatic lines of force are unable to crowd into these pockets and pull the electrons away as fast as they are emitted, and the cloud of electrons which thereby collects in these cavities forms a space charge that repels the further electrons which are emitted and tends to drive them back to the cathode. Consequently, any method of heat shielding designed to reduce the loss by radiation will also impede the migration of electrons. Only when positive ions are present and free to move into the cavity and neutralize the space charge will the area lining the cavity emit all the electrons it would if it were fully exposed. In the mercury-vapor tubes using the above form of cathode construction a continuous supply of positive ions is formed by the collision of fast-moving electrons with the vapor particles, and these ions neutralize the space charge so that the migration of electrons is unimpeded.

This conservation of the filament heat requires appreciable time before the emitting surface reaches the proper operating temperature. The time required ranges from 30 sec. to 30 min., depending on the size of the tube; the latter figure being for a mercury-vapor rectifier having a total emission of over 450 amp. and requiring 330 watts for cathode heating.

The low-temperature emitters referred to above consist of a coating of the oxides of barium and strontium on a metal core. This discovery was made by Wehnelt[8] in 1904 and is known as an *oxide-coated cathode*. Formerly the core material used was platinum, but its expense led to a search for other materials.

[8] *Ann. Physik*, vol. 14, p. 425, 1904.

It appears that the core material may play some part in the emission process and is not merely a supporting electrical conductor. Alloys of nickel and platinum are still used and give very long life, 20,000 hr. being common with certain types of tubes. The more usual material at the present time is pure nickel; or alloys of nickel, such as "Konel" which contains nickel, cobalt, iron, and titanium. Cathodes of this type produce a relatively large emission at low temperatures. Most of the power required to maintain the cathode at a given temperature is represented by the energy radiated as heat. The small loss by conduction is along the support and lead wires, since the cathode is in a high vacuum. The energy radiated per unit area varies as the fourth power of the absolute temperature so that an emitter such as tungsten which is normally operated at a temperature of 2400°K. will require considerably more cathode heating power than oxide-coated cathodes of the same area whose operating temperatures range from 900 to 1200°K. The melting point of tungsten is 3655°K.

The active material is applied to the ribbon-shaped filament by passing it through a thick aqueous suspension of very finely divided barium and strontium carbonates. Several applications are usually necessary to obtain a coating of the desired thickness. The filament is passed through a drying oven after each application. In heater-type cathodes the carbonates are frequently applied by a spray-gun. While the tube is being exhausted, the cathode is heated to several hundred degrees above the normal operating temperature which converts the carbonates into oxides with the evolution of carbon dioxide. The high temperature also serves to activate the cathode. This activation process, which is more of an art than a science, is sometimes supplemented by bombarding the cathode with positive ions by impressing a positive potential of 100 to 200 volts on the plate before the tube is highly evacuated.

The copious emission of electrons in oxide-coated filaments seems to be from a layer of metallic barium only one molecule thick which is formed on the surface of the oxide.[9]

A third form of emitter is the *thoriated-tungsten filament.* These filaments consist of tungsten containing a few per cent of

[9] J. A. BECKER, *Bell Lab. Rec.*, vol. 9, p. 54, October, 1930; or *Electronics*, vol. 1. p. 390, November, 1930.

thorium oxide, together with some reducing agent such as carbon. They are activated after the tube has been thoroughly exhausted by heating to a temperature of 2600 to 2800°K. for a minute or more, after which the temperature is reduced to a value of 300 or 400° above the normal operating temperature of 1900°K. and held there for some minutes. The action of the high temperature causes some of the thorium oxide to be reduced to metallic thorium by action of the reducing agent and the subsequent heating diffuses this thorium to the surface in much the same manner as the barium layer in the oxide-coated filament. As the thorium evaporates from the surface with use, it is replenished by further diffusion from the interior of the tungsten.

This type of filament, while not having an emission efficiency so high as the oxide-coated type, is more rugged and may be used with plate voltages which would cause the oxide-coated type to disintegrate too rapidly. Thoriated filaments which have lost their emission can frequently be reactivated, provided the thorium content of the filament is not exhausted. The process consists of impressing 350 per cent normal voltage across the filament of the tube for about 15 sec., after which the filament voltage is reduced to about 150 per cent normal for an hour or two with the plate voltage removed. In the newer carbonized types of filament the initial high voltage application is omitted. If the loss of emission has been due to a poor vacuum, it is useless to attempt reactivation. The film of thorium is chemically very active and tubes employing these filaments must be very thoroughly evacuated before they are activated. Traces of oxygen and water vapor are particularly detrimental. Positive-ion bombardment will also destroy the surface film of thorium so that a high vacuum is essential. This is secured by vaporizing within the tube some chemically active substance such as magnesium to absorb the residual gas just before the tube is sealed off from the vacuum pump.

A recent improvement in thoriated filaments is to carbonize the surface of the tungsten. This can be done by heating the filament above 1600°K. in a hydrocarbon vapor at low pressure. The vapor molecules which strike the hot surface decompose into carbon and hydrogen. The carbon diffuses into the tungsten forming a shell of tungsten carbide on the outer surface. The evaporation of thorium from the carbonized surface of the fila-

ment is greatly reduced so that higher operating temperatures can be used without deactivating the filament. At a temperature of 2200°K. the evaporation of thorium is only about one-sixth of that from a tungsten surface of the same temperature. The higher working temperatures permissible with this type of thoriated filament greatly increase the emission that can be obtained from a given area of filament. The carbonized filament also gives the thorium surface greater immunity from damage due to positive-ion bombardment.

49. Comparison of Emitters.—The comparative emission efficiencies of the three types of emitters are shown in Fig. 78. The saturation current can be expressed with a fair degree of accuracy in the form

$$I_s = cP^n \tag{6}$$

where c and n are constants for a particular filament and P is the heating power supplied. Plotting I_s against P on logarithmic paper results in a slight departure from a straight-line relationship. If the coordinates of the logarithmic paper are slightly distorted, as is done in Fig. 78, a straight line will result. Graph paper ruled in this fashion is known as power-emission paper. Only two points are needed to obtain the complete performance of a given cathode, which, in addition to saving time, permits the observations to be made at temperatures lower than the normal value. This latter item is of considerable importance, as the plate voltage required to produce complete saturation at normal filament temperature will be excessively large and may ruin the tube by reason of excessive plate heating.

The temperatures at the top of the figure apply only to tungsten and thoriated tungsten. Oxide-coated cathodes vary considerably as to color and roughness of the surface, type of core material, etc., all of which affects the radiation so that a temperature scale would apply to one particular type only.

As will be seen from Fig. 78, the oxide-coated cathode has the highest emission efficiency. Commercial receiving tubes of the filament type will ordinarily average about 600 milliamperes per watt. The efficiency of the heater type is much lower, ranging from 30 to 100 milliamperes per watt, which is about the same as the thoriated-tungsten filaments. Pure tungsten will vary from 2 to 10 milliamperes, depending on the life desired. Longer life

is secured by operating the filament at lower temperatures and thus reducing the rate of filament evaporation, which, of course, reduces the emission efficiency. Thoriated filaments tend to

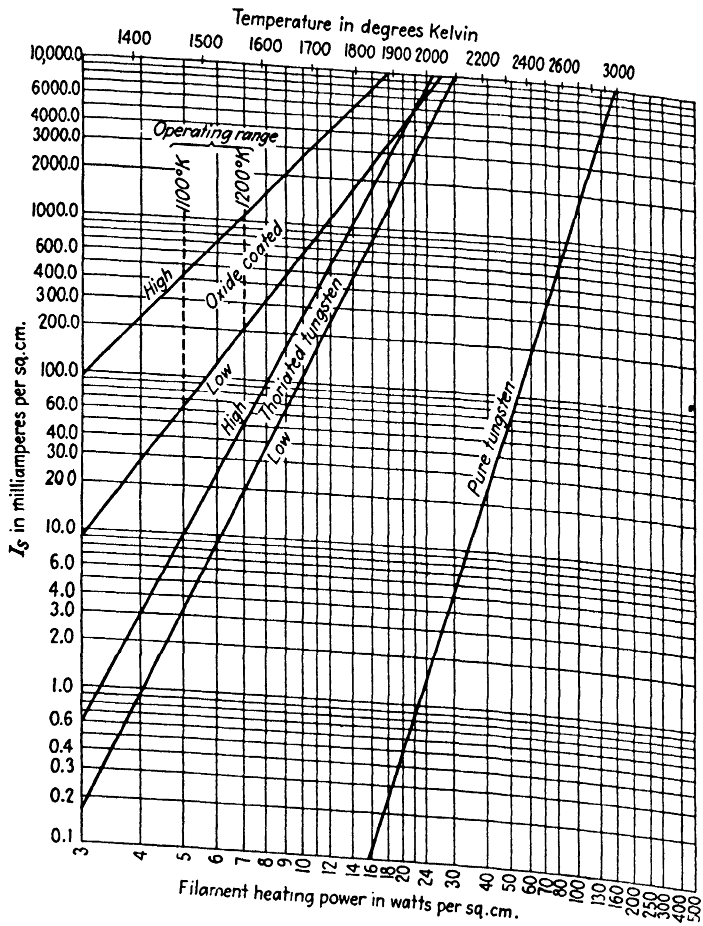


FIG. 78.—Comparative emission efficiencies showing the relation between cathode heating power and total emission for three types of cathodes. The temperature scale applies only to tungsten and thoriated tungsten.

lose their emission if operated below normal temperature with the plate voltage at rated value, as the diffusion of thorium to the filament surface will not be rapid enough to make up for the loss by evaporation. This is particularly true in the case of power tubes.

Oxide-coated filaments usually have a longer life than the other two types, particularly in the smaller tubes. Heater-type cathodes are always oxide-coated because of the low temperatures at which adequate emission can be obtained. It is more difficult to secure as high a vacuum which prevents the use of oxide-coated filaments in the larger power tubes as these filaments are readily injured by positive-ion bombardment. In this field the thoriated filament is very satisfactory. Tungsten filaments have to be employed in all of the high-power tubes as the ordinary thoriated type is unable to withstand the very high plate voltages used and the active layer of thorium is stripped off faster than internal diffusion can replace it.

50. Effects of Gas.—As has been mentioned previously, electrons traveling with sufficient velocity may, upon striking a gas molecule, knock out one or more of its constituent electrons. This has removed one or more negative charges from the gas molecule, leaving it positively charged, or ionized. This process is known as ionization by collision. The velocity needed by an electron to ionize a gas in this manner is different for each gas and is usually expressed in terms of the potential difference through which an electron must fall in order to acquire the necessary velocity. This potential difference is called the *ionizing potential* of the gas. The ionizing potentials of most gases range from 10 to 25 volts. The positive ions produced in this way in a vacuum tube are repelled by the plate and are attracted to the negatively charged cathode. The mass of these ions is very much greater than that of an electron—about 1840 times as great in the case of hydrogen—and if they are produced in sufficient numbers and are subjected to a strong electrostatic field, their bombardment of the cathode will cause it to disintegrate rapidly. The cathode is nearly always subjected to a small amount of positive-ion bombardment, even in well-evacuated tubes, as it is impossible to remove all traces of residual gas.

The presence of these positive ions serves to annul the space charge produced by the electrons so that by using the proper gas at a suitable pressure it is possible to obtain saturation current at comparatively low values of plate potential. This principle is employed in the Tungar rectifier which uses argon at a pressure of about 1 lb. absolute. With this device saturation currents of

several amperes can be obtained with an internal drop in the tube of only 5 to 10 volts. A heavy tungsten filament is used to withstand the positive-ion bombardment, which is of reduced intensity because of this low voltage drop. The filament temperature is higher than that used in high-vacuum tubes, causing the emission to be about ten times as great as ordinarily obtained from tungsten. This is feasible only because the gas pressure within the tube is sufficient to reduce the rate of filament evaporation to a nominal amount. The effect of positive-ion bombardment can be shown in an interesting fashion in these tubes by disconnecting one side of the filament heating circuit while the device is in operation. The bombardment is sufficient to keep the opposite end of the filament white hot so that the tube continues to function.

The neutralization of the space charge by positive ions is also used in the hot-cathode mercury-vapor rectifiers, which have already been mentioned in connection with heat-shielded cathodes. These cathodes are oxide-coated and withstand the effects of positive-ion bombardment only by virtue of the low internal drop in the tube. It has been found that if this drop is kept below 25 volts, the velocity of the ions is not sufficient to injure the emitting surface.[10] While these tubes are used in circuits rectifying thousands of volts, it must be remembered that nearly all of this potential is consumed across the useful load and that the drop across the tube is only from 10 to 15 volts. It is exceedingly important that the cathodes of these tubes be allowed to reach operating temperature before the plate voltage is applied. If the plate voltage is applied before the electron emission has reached normal value, the internal drop in the tube will be excessive and the cathode surface will be ruined. This is taken care of in commercial applications by the use of suitable time-delay relays.

A small quantity of metallic mercury is present in the evacuated bulb so that the mercury-vapor pressure is governed by the temperature of the coolest spot within the tube where condensation of the vapor takes place. Consequently, due consideration must be given to the operating temperature of these tubes and adequate ventilation must be supplied, as too high a temperature will reduce the inverse or "flash-back" voltage that can be

[10] A. W. HULL, *Trans. A.I.E.E.*, vol. 47, p. 753, 1928.

withstood, while too low a temperature will increase the internal drop to the point where the cathode may be injured.

In the high-vacuum types of thermionic tubes the presence of gas is highly objectionable because of the much larger internal voltage drops present which would give rise to destructive bombardment of the cathode by the positive ions. The presence of appreciable gas also causes erratic variations in the plate-current characteristic. This was somewhat of an advantage when the vacuum tube was first introduced, as these irregularities in the characteristic greatly increased the sensitivity of the device as a detector. The first de Forest audions were all poorly evacuated or "soft" and would produce the characteristic bluish glow of gaseous ionization, even with relatively low plate voltages. The plate voltage and filament temperature had to be critically adjusted for maximum sensitivity. As the technique of radio-frequency amplification developed, the need for high sensitivity on the part of the detector subsided and soft tubes as detectors fell into disuse.

Gaseous types of cold-cathode rectifier tubes[11] were formerly used to a considerable extent in receiving sets and B battery eliminators, but have since been practically all replaced by the hot-cathode type.

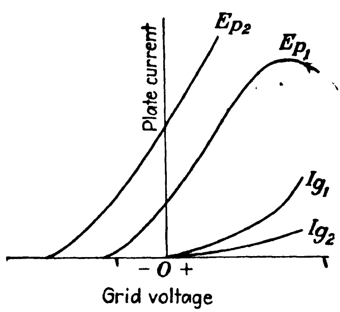


FIG. 79.—Variation of plate and grid currents with grid voltage in a triode.

51. Triodes.—The introduction of the third electrode or grid between the plate and cathode offers a means of controlling the flow of electrons to the plate. By making the potential of the grid negative with respect to the cathode the effect of the space charge can be augmented, decreasing the plate current.

If the grid potential is sufficiently negative, the plate current will be reduced to zero. The value of grid voltage required to accomplish this is called the *cut-off voltage*.

A positive charge on the grid partially neutralizes the space charge, which increases the plate current. With the grid positive, some of the electrons are attracted to it, causing a current to

[11] V. BUSH and C. G. SMITH, *Jour. A.I.E.E.*, p. 627, September, 1922.

flow into the grid. As the grid is made increasingly positive, the grid current continues to rise, and since the total number of electrons emitted by the cathode is fixed by the cathode temperature, the plate current ultimately begins to fall as shown in Fig. 79, particularly at low plate voltages. If the plate voltage is increased to E_{p2} the competition offered by the grid is correspondingly less, resulting in a lower grid current and a higher saturation value of plate current.

Since the grid is located closer to the cathode than the plate is, a small change in the grid potential can offset a very much larger change in the plate potential. If an alternating voltage is applied to the grid it is then equivalent to the introduction of a much larger alternating voltage inserted in the plate circuit. This enables the tube to function as an amplifier.

52. Characteristic Curves of Triodes.—The plate current is a function of both the grid and plate voltages, or expressed mathematically

$$I_p = f(E_g, E_p) \tag{7}$$

It also depends on the filament current, but as this is usually kept constant at rated value, the effect of variations in this item will not be taken into account. The shape of the plate current curves of Fig. 79 is similar to the diode of Fig. 76, and variations in filament temperature will vary the saturation current of a triode in a like manner.

Figure 80 shows a family of characteristic curves illustrating the variation of plate current with grid voltage for various values of plate voltage in the case of a typical triode of the heater type. This group will be referred to as the I_p-E_g characteristics. The only effect of increasing the plate voltage is to shift the curves to the left without changing their slope. If the plate voltage is made the independent variable, as in Fig. 81, another very useful group of curves is obtained, known as the I_p-E_p characteristics. The tops of both groups of curves will ultimately bend over and become horizontal as saturation is approached.

If the plate and grid voltages are simultaneously varied so as to keep the plate current constant, a third group of characteristics is obtained, as shown in Fig. 82. These show the relative effects of the grid and plate voltages on the plate current of the

tube and are of interest in that their slope defines one of the important tube constants.

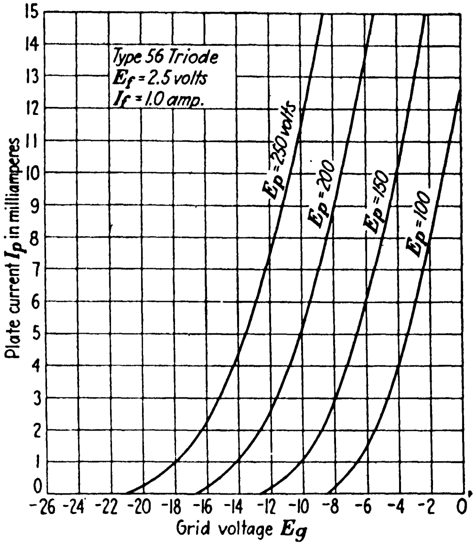


FIG. 80.—Plate-current grid-voltage characteristics for a typical triode.

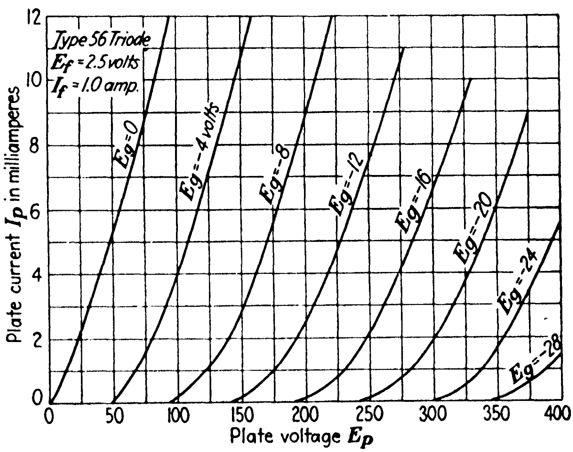


FIG. 81.—Plate-current plate-voltage characteristics of a typical triode.

53. Triode Constants.—The *amplification factor* μ of a triode is defined as the ratio of the change in plate voltage to a change in grid voltage, the plate current remaining unchanged. The grid

voltage change is seen to be in a direction opposite to the change in plate voltage, *i.e.*, a decrease in plate voltage is offset by an increase in the grid voltage. Mathematically, this definition becomes

$$\mu = -\frac{\partial e_p}{\partial e_g}, \qquad di_p = 0 \tag{8}$$

and is equal to the slope of the curve in Fig. 82, which is negative

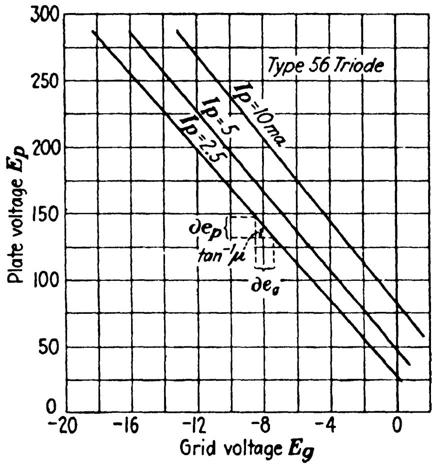


FIG. 82.—Plate-voltage grid-voltage characteristics. The amplification factor μ is equal to the slope of the curve.

in the mathematical sense. Partial derivatives are used since the third variable i_p is held constant. This is implied by $di_p = 0$.

The slope at any point on the curves of Fig. 80 will be equal to an increment of plate current divided by a corresponding increment of grid voltage and hence is of the nature of a conductance. This has been called the *mutual conductance* and is defined as

$$g_m = \frac{\partial i_p}{\partial e_g}, \qquad de_p = 0 \tag{9}$$

Recently the name *grid-plate transconductance* is being used to an increasing extent for this quantity, using the symbol s_m, or, more accurately, s_{pg}, so as to fit in with the nomenclature which had to be developed to take care of multi-grid tubes. *Transconductance* is defined as the ratio of a change in current in the circuit of one electrode to the change in potential on another electrode, all other voltages remaining unchanged; or mathematically

$$s_{jk} = \frac{\partial i_j}{\partial e_k}, \qquad de = 0 \tag{10}$$

where j and k are the two electrodes in question.

The slope at a point on the I_p-E_p characteristic will be the *plate conductance* g_p. The reciprocal of this quantity is the *plate resistance* which is defined as

$$r_p = \frac{1}{g_p} = \frac{\partial e_p}{\partial i_p}, \qquad de_g = 0 \tag{11}$$

and is the apparent resistance offered by the plate circuit to a small increment of plate voltage. This is not the same as the resistance offered to the battery in the plate circuit as will be seen from Fig. 83. The resistance encountered by the battery is

$$R_p = \frac{E_p}{I_p} = \cot \theta \tag{12}$$

while that offered to an incremental change, such as a small alternating voltage superimposed on E_p, is

$$r_p = \frac{\Delta e_p}{\Delta i_p} = \cot \phi$$

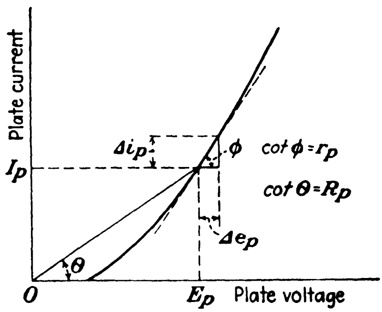


FIG. 83.—Graphical determination of static and dynamic plate resistances of a triode.

These two resistances will be the same only if the characteristic is a straight line passing through the origin. Equation (11) may be thought of as the *dynamic plate resistance* while (12) is the *static plate resistance.* The former is the one with which we are chiefly concerned and is the value implied whenever the plate resistance is referred to, unless otherwise stated.

The variations in these "constants" are shown in Fig. 84, for a particular value of plate voltage. The values will be slightly different for other plate voltages, but the general trend will be the same. The amplification factor μ remains reasonably constant for a relatively wide range of conditions. The plate resistance r_p will be infinite at cut-off voltage and it again approaches infinity in the vicinity of saturation, where the slope of the I_p-E_p characteristic approaches zero. For a small variation of plate

current, r_p may be regarded as a constant with sufficient accuracy for approximate calculations.

The cut-off voltage E_{co} of a tube is defined as the value of negative grid voltage required to reduce the plate current to zero

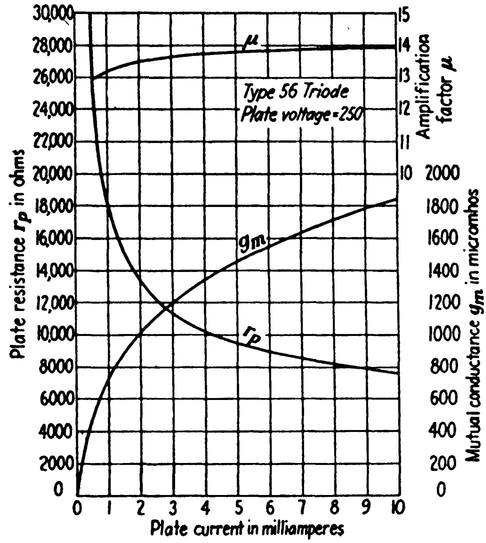


FIG. 84.—Variation of amplification factor, mutual conductance, and plate resistance with plate current for a fixed value of plate voltage. Plate current varied by varying grid voltage.

and is given by the approximate relation

$$E_{co} = \frac{E_p}{\mu} \tag{13}$$

The relationship existing among these various tube constants is

$$\mu = \frac{\Delta e_p}{\Delta e_g} = \frac{\Delta i_p}{\Delta e_g} \cdot \frac{\Delta e_p}{\Delta i_p} = g_m r_p \tag{14}$$

so that if any two are known the third can be determined.

54. Expressions for the Plate Current.—The plate current of a triode is a function of both the grid and plate voltages and may be expressed by

$$I_p = K(E_p + \mu E_g)^{3/2} \tag{15}$$

where K is a constant related to the plate conductance. The quantity in parenthesis is termed the *equivalent plate voltage.*

The expression assumes that the current follows the theoretical three-halves power law. In actual practice the value of the exponent varies from a trifle over unity to about 2.5.

The current can also be expressed as

$$I_p = K'\left(E_g + \frac{E_p}{\mu}\right)^{3/2} \tag{16}$$

where K' is a constant related to the mutual conductance g_m. The quantity in parentheses in this case is called the *equivalent grid voltage*. For small variations of grid voltage the characteristic may be assumed to be linear and in this region the plate current is given by

$$I_p = \frac{1}{r_p}(E_p + \mu E_g - e_0) \tag{17}$$

where e_0 is the intercept of the I_p-E_p characteristic with the plate voltage axis.

When the value of the exponent in the above equations is not known, or if it varies appreciably over the operating range, an analytical solution may be had by representing the characteristic in the form of a Taylor's series. This method of analysis may be applied to any circuit containing a nonlinear impedance, which is a circuit or device which does not follow Ohm's law. In addition to vacuum tubes, other examples of nonlinear impedances are contact rectifiers of various types, saturated iron-core reactors, and certain nonmetallic conductors.

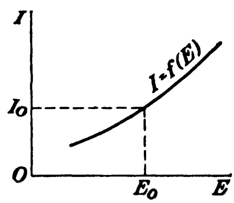


FIG. 85.—Voltage and current relations in a nonlinear impedance.

Assume that the characteristic of the nonlinear impedance is represented by Fig. 85, and that we are operating about the point (E_0, I_0) by superimposing a small alternating voltage on E_0. The characteristic can be expressed as a power series of the form

$$I = I_0 + a(E - E_0) + b(E - E_0)^2 + c(E - E_0)^3 + d(E - E_0)^4 + \cdots \tag{18}$$

As many terms can be used as the accuracy of the solution demands.

To evaluate the constants, differentiate (18) with respect to E, which gives us

$$\frac{dI}{dE} = a + 2b(E - E_0) + 3c(E - E_0)^2 + 4d(E - E_0)^3 + \cdots \quad (19)$$

The derivative dI/dE is to be evaluated at the point (E_0, I_0) and is equal to the slope of the characteristic at this point. Since at this point $E - E_0 = 0$, (19) becomes

$$\left[\frac{dI}{dE}\right]_{E_0, I_0} = a \quad (20)$$

and has the dimensions of a conductance.

The constant b is found by differentiating (19) with respect to E, giving

$$\frac{d^2I}{dE^2} = 2b + 6c(E - E_0) + 12d(E - E_0)^2 + \cdots \quad (21)$$

which, when evaluated at the point (E_0, I_0), becomes

$$\frac{1}{2}\left[\frac{d^2I}{dE^2}\right]_{E_0, I_0} = b \quad (22)$$

The nth constant will be

$$\frac{1}{\underline{|n}}\left[\frac{d^nI}{dE^n}\right]_{E_0, I_0} = k_n \quad (23)$$

Substituting these values in (18)

$$I = I_0 + \left[\frac{dI}{dE}\right]_{E_0, I_0}(E - E_0) + \frac{1}{2}\left[\frac{d^2I}{dE^2}\right]_{E_0, I_0}(E - E_0)^2 + \frac{1}{6}\left[\frac{d^3I}{dE^3}\right]_{E_0, I_0}(E - E_0)^3 + \cdots + \frac{1}{\underline{|n}}\left[\frac{d^nI}{dE^n}\right]_{E_0, I_0}(E - E_0)^n \quad (24)$$

In the case of a triode either the equivalent grid voltage or the equivalent plate voltage can be substituted for E in (24). The application of Taylor's series to a triode will be given detailed consideration in Chap. XI.

55. Vacuum-tube Notation.—In order to avoid confusion, the following notation will be used so far as possible to designate the various currents and voltages in the circuits of the several electrodes. The subscripts a, b, and c will be used to denote quantities in the *external* filament, plate, and grid circuits, respectively. The subscripts f, p, and g will denote *internal*, or

tube, quantities in filament, plate, and grid. Capital letters will be used whenever possible for constant or direct-current values, while small letters will be employed for alternating-current values.

An illustration of this system is given in Fig. 86. The polarity of the C battery in the grid circuit is purposely reversed so as to

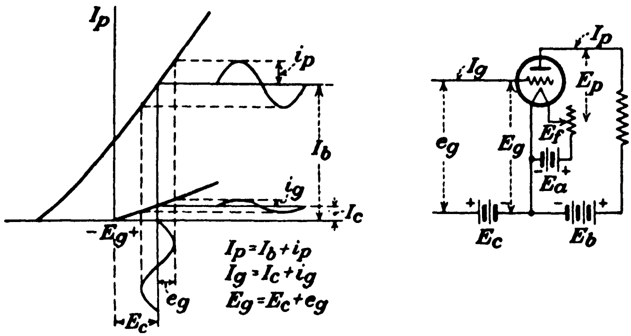


FIG. 86.—Illustration of vacuum-tube notation used.

bias the grid with a positive potential which will cause grid current to flow. This is merely for illustration, as the normal function of the C battery is to hold the grid sufficiently negative so that no grid current flows. The voltage E_c is termed the *grid bias.*

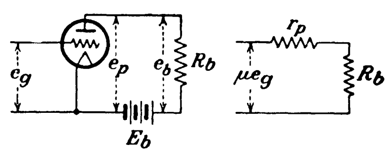


FIG. 87.—Actual and equivalent circuits of a triode.

56. Equivalent Circuit of a Triode.—If a small alternating voltage e_g is applied to the grid of a triode, as in Fig. 87, the alternating component of plate current may be determined by assuming that the tube has an e.m.f. μe_g acting in the plate circuit through an internal resistance r_p. The alternating plate current will be some function of the alternating grid and plate voltages, which is expressed mathematically by

$$i_p = f(e_p, e_g) \tag{25}$$

The total differential of this function is

$$\begin{aligned} di_p &= \frac{\partial i_p}{\partial e_p} de_p + \frac{\partial i_p}{\partial e_g} de_g \\ &= g_p de_p + g_m de_g \end{aligned} \tag{26}$$

Since $\mu = r_p g_m$, (26) becomes

$$di_p = \frac{1}{r_p}(de_p + \mu de_g)$$

or

$$r_p = \frac{de_p}{di_p} + \mu\frac{de_g}{di_p} \tag{27}$$

From Kirchhoff's laws

$$E_b - e_b - e_p = 0$$

Differentiating,

$$dE_b - de_b - de_p = 0$$

But as E_b is a constant, $dE_b = 0$, so that

$$de_p = -de_b \tag{28}$$

Substituting this value for de_p in (27), we get

$$r_p = -\frac{de_b}{di_p} + \mu\frac{de_g}{di_p} = -R_b + \mu\frac{de_g}{di_p}$$

or

$$di_p = \frac{\mu de_g}{r_p + R_b} \tag{29}$$

Integrating both sides,

$$i_p = \frac{\mu e_g}{r_p + R_b} \tag{30}$$

which is evidently the expression for the current in the equivalent circuit of Fig. 87. Therefore a triode circuit can be handled as though it were an alternator having an e.m.f. μe_g and an internal resistance r_p. If the load in the plate circuit had been an impedance $Z_b = R_b + jX_b$, instead of a resistance, the current would then have been

$$i_p = \frac{\mu e_g}{r_p + R_b + jX_b} \tag{31}$$

The scalar magnitude would be

$$|i_p| = \frac{\mu e_g}{\sqrt{(r_p + R_b)^2 + X_b^2}} \tag{32}$$

These expressions assume that r_p and μ remain constant as e_g varies, which is approximately true if the amplitude of e_g is small, or if the characteristic is made linear. For larger values

of grid voltage, such as are impressed upon the grid of power-amplifier tubes in a receiving set, r_p may vary appreciably throughout the cycle. In this case we must either resort to a graphical solution or else use Taylor's series. Equation (30) is the second term in the series of (24) in the case of a pure resistance load in the plate. The first term is the direct-current component of plate current I_b. The third, fourth, and higher order terms of the series represent the second, third, and higher order harmonics of the plate current, assuming the grid voltage to be given by $e_g = E_m \sin \omega t$. These higher order terms represent distortion in the case of an amplifier, so that (30) gives the amplitude of the fundamental component of the alternating plate current. In the case of certain types of detectors and modulators the third term of (24) is the useful one, since rectification is a distortion process.

The equivalent circuit of Fig. 87 also neglects the effects of capacitance between the various electrodes. These are of the order of a few micro-microfarads in the ordinary receiving tube and can be usually neglected except for high values of frequency.

57. Calculation of Triode Constants from the Structural Dimensions.—The constants of a triode having plane electrodes may be determined to a fair degree of accuracy from the following expressions due to R. W. King.[12]

$$\mu = \frac{2\pi b_2 n}{\log_\epsilon \dfrac{1}{2\pi\rho n}} \tag{33}$$

$$K = \frac{2.336 \times 10^{-6} A}{(b_1 + b_2)^{1/2}[b_2 + b_1(\mu + 1)]^{3/2}} \tag{34}$$

where b_1 = distance from filament to grid.
b_2 = distance from grid to plate.
ρ = radius of grid wires.
n = number of grid wires per centimeter.
A = effective area of plate.

all dimensions being in centimeters. Equation (34) gives the value of the constant K in (15).

[12] *Phys. Rev.*, vol. 15, p. 256, 1920. For a comprehensive treatment of this subject the reader is referred to a paper by QUZIRO KUSUNOSE, Calculation of Characteristics and the Design of Triodes, *Proc. I.R.E.*, vol. 17, p. 1706, October, 1929.

If the plate and grid are cylinders,

$$\mu = \frac{2\pi\rho_g^2\left(\dfrac{1}{\rho_g} - \dfrac{1}{\rho_p}\right)n}{\log_e \dfrac{1}{2\pi\rho n}} \tag{35}$$

$$K = \frac{14.68 \times 10^{-6} l \sqrt{\rho_p}}{\left(\rho_g + \dfrac{\rho_p}{\mu}\right)^{3/2}} \tag{36}$$

where ρ_g and ρ_p are the radii of the grid and plate, and l is the length of the plate, all dimensions in centimeters.

In the case of a tube with plane electrodes the value of μ increases directly with the number of grid wires per unit of length and with the distance from grid to plate. With cylindrical electrodes, if all dimensions are constant except the radius of the grid, (35) may be written in the form

$$\mu = C\left(\rho_g - \frac{\rho_g^2}{\rho_p}\right)$$

Differentiating this with respect to ρ_g and setting the resultant expression equal to zero, we get

$$\frac{d\mu}{d\rho_g} = C\left(1 - \frac{2\rho_g}{\rho_p}\right) = 0$$

or

$$\rho_g = \tfrac{1}{2}\rho_p \tag{37}$$

so that μ is a maximum in a triode with cylindrical electrodes when the grid is placed midway between the cathode and plate.

58. Measurement of Triode Constants.—The values of μ, r_p, and g_m may be determined graphically from the characteristic curves of the tube by drawing tangents to the curves of Figs. 80, 81, and 82 and determining the slope at the points in question. The accuracy obtained in this manner is not very great and it is usually more desirable to measure these quantities dynamically by means of a suitable bridge.

The amplification factor is readily determined by the circuit of Fig. 88. The value of R_1 is about 10 ohms and R_2 is adjusted for no sound in the telephone receivers. The value of μ is given by

$$\mu = \frac{R_2}{R_1} \tag{38}$$

The condenser C is sometimes necessary to balance out the internal tube capacitances and secure a good null point. The telephone receivers are preferably connected to the secondary of a small step-up transformer offering low primary resistance to the flow of the direct-current component of plate current. Shunting

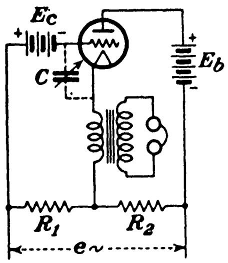


FIG. 88.—Circuit for the measurement of μ.

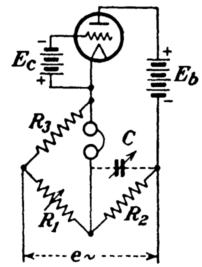


FIG. 89.—Circuit for the measurement of r_p.

the telephone receivers across a low resistance choke coil is an alternate method. The impressed alternating voltage e should be no larger than is necessary to secure a good balance.

The plate resistance r_p can be measured by using the plate circuit of the tube as the fourth arm of an ordinary bridge as shown in Fig. 89. When the bridge is balanced,

$$r_p = \frac{R_2R_3}{R_1} \tag{39}$$

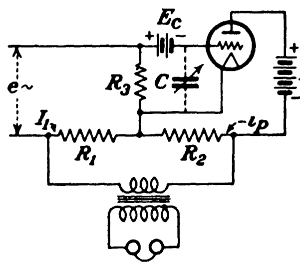


FIG. 90.—Circuit for the measurement of g_m.

R_3 is about 10,000 ohms, or some value comparable to r_p, and R_2 is fixed at 10 or 100 ohms. A balance is secured by varying R_1, which will be comparable to R_2 in magnitude. The variable condenser C balances out tube capacitances as before. If r_p is large, the value of C needed may be inconveniently large, in which case C may be shunted across R_3. A choke coil or transformer should be used with the telephone receivers, as in the measurement of μ.

The mutual conductance can be measured by the arrangement of Fig. 90. For a balance

$$I_1R_1 = i_pR_2 = \frac{\mu e_g}{r_p + R_2}R_2 = \frac{\mu I_1R_3}{r_p + R_2}R_2$$

If R_2 can be neglected in comparison with r_p, the above expression becomes

$$R_1 = \frac{\mu}{r_p}R_3R_2 = g_mR_3R_2$$

or

$$g_m = \frac{R_1}{R_2R_3} \tag{40}$$

The variable condenser C may be used to balance internal capacitances, if necessary. The resistance R_1 can be varied to secure a balance, while R_3 may be a fixed value of 1000 ohms and R_2 fixed at 100 ohms. The mutual conductance is usually expressed in micromhos, so that with the above values of resistance, g_m in these units will be ten times the setting of R_1.

For an excellent discussion of the various dynamic methods of measuring triode constants the reader is referred to "Theory of Thermionic Vacuum Tubes,"[13] Chap. IX, by E. L. Chaffee.

Problems

1. The tungsten filament of a power tube is 5.15 in. long and has a diameter of 16 mils (0.016 in.). The operating temperature is 2500°K. at which value the specific resistance of tungsten is 444.5 ohms per circular mil-foot. What is the total saturation current? What is the filament current if the potential across the filament is 11 volts?

2. The filament in the above tube is replaced with a thoriated filament having the following emission constants: $A = 5$, $b_0 = 31{,}500$. The operating temperature of the new filament is 1900°K. and the specific resistance at this temperature is 321 ohms per circular mil-foot. The power required to maintain this temperature is 18.6 watts per square centimeter. If the filament voltage and total saturation current are to be the same as before, find the length and diameter of the new filament. What will be the filament current?

3. The characteristic of a triode is given by

$$I_p = 0.0021(E_p + \mu E_g)^2 \text{ milliamperes}$$

where $\mu = 13.8$. It is operated at a plate potential of 200 volts and a negative grid bias of -12 volts. If a signal voltage of $e_g = 2 \sin \omega t$ volts is impressed on the grid, find the maximum, minimum, and average values of plate current. What is the percentage of second harmonic current in the plate circuit? What is the plate current when $e_g = 0$?

[13] McGraw-Hill Book Company, Inc., 1933.

4. A triode having plane electrodes has the following structural dimensions in centimeters: $b_1 = 0.169$, $b_2 = 0.466$, $1/n = 0.29$, $\rho = 0.01$, $A = 12$. What is the amplification factor? What is the plate current for a plate potential of 100 volts and a negative grid bias of -10 volts, assuming the plate current follows the three-halves power law? What is r_p?

5. Design a plane electrode tube having $\mu = 10$ and $r_p = 10{,}000$ ohms at $E_p = 500$ volts and $E_c = -5$ volts.

6. A choke coil in the plate circuit of a tube having an amplification factor of 9 and a plate resistance of 9000 ohms, has an inductance of 10 henrys and a value of $Q = 8$. What will be the alternating voltage across the choke coil if a 60-cycle potential of 5 volts is impressed across the grid of the tube?

7. The choke coil of Problem 6 is shunted by a condenser so as to make the combination a parallel-resonant circuit. What will be the alternating voltage across the parallel circuit? What is the value of capacitance needed?

8. A variable resistance R_b is inserted in the plate circuit of the tube in Problem 6 in place of the choke coil. What will be the maximum alternating-current power that the tube can supply to R_b if $e_g = 5$ volts? What is the value of R_b for maximum power?

CHAPTER VII

AUDIO-FREQUENCY AMPLIFIERS

59. Vacuum-tube Amplifiers and Their Classification.—The amplifying action of a vacuum tube is due to the fact that a small voltage applied to the grid is the equivalent of a much larger voltage acting in the plate circuit of the tube. Almost any amount of amplification can be secured by employing a number of tubes in cascade, the amplified output voltage of one tube serving as the input voltage of the next.

The basic circuit of an amplifier is shown in Fig. 91. The alternating component of plate current i_p flows through the load

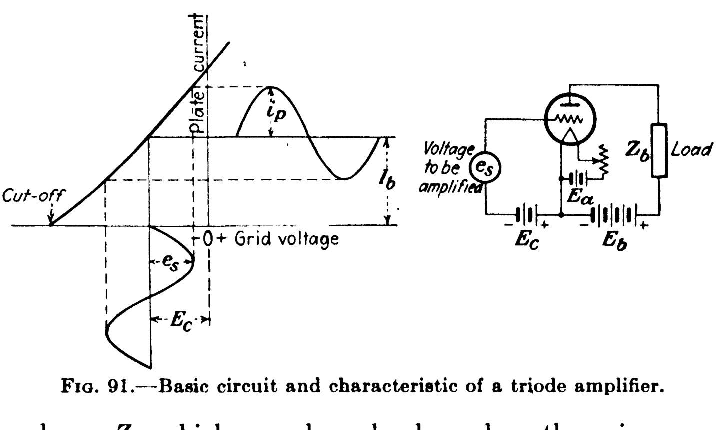


FIG. 91.—Basic circuit and characteristic of a triode amplifier.

impedance Z_b, which may be a loud-speaker, the primary of a transformer, or any other device. The voltage supplied to the load will be i_pZ_b, and the *voltage amplification* will be

$$A_v = \frac{i_pZ_b}{e_s} \tag{1}$$

where e_s is the signal voltage to be amplified. There will also be a voltage drop across the load due to the direct-current component of plate current I_b, equal to I_bR_b, where R_b is the resistance offered by the load to the flow of direct current. The purpose

of the voltage E_c is to bias the grid sufficiently negative so that the instantaneous grid potential never becomes positive with respect to the cathode. If the grid were allowed to become positive, it would attract electrons and a current would flow in the grid circuit. The energy represented by this current has to be supplied by the source of e_s, which is usually objectionable.

Amplifiers may be classified either as to their frequency range or as to their mode of operation. The former classification divides them into audio-frequency, radio-frequency, and direct-current amplifiers. Each of these groups may be further subdivided as to the nature of the coupling between stages and other details of operation. The classification as to the mode of operation depends upon how large a portion of the characteristic is being used. These groups are as follows:

Class A amplifiers are those operated so that the output-wave shapes of the plate current are practically the same as those of the exciting grid voltage. The grid must not go positive on excitation peaks and the plate current must not fall low enough at its minimum to cause distortion due to the curvature of the characteristic in this region. Most of the amplifiers used in radio reception operate as Class A devices.

Class B amplifiers are operated with a negative grid bias approximately equal to cut-off so that the plate current is almost zero when the grid excitation is removed. With a sinusoidal voltage applied to the grid, the plate current consists of a series of half-sine waves, similar to the output of a half-wave rectifier. The grid usually swings positive on excitation peaks, causing grid current to flow. Class B amplifiers are used in radiotelephone transmitters following the modulated stage. They are also being employed as audio-frequency amplifiers using two tubes which operate so that one tube amplifies the positive half cycle of the signal voltage while the other tube amplifies the negative half cycle. Class B operation is characterized by larger power output and higher efficiency than Class A.

A *Class C* amplifier is one in which high output is the primary consideration and is employed only in radio transmitters. The grid is negatively biased to a point considerably beyond cut-off so that the plate current is zero with no grid excitation. The grid-excitation voltage used is large and is often sufficient to cause the plate current to reach saturation on the positive swings,

resulting in a large grid current. The power output and efficiency are both higher than with Class B operation.

Other intermediate modes of operation, such as Class AB, will be discussed later.

60. Distortion in Amplifiers.—An ideal Class A amplifier would produce a plate-current-wave shape that would be identical in form with the impressed grid-voltage wave. Any distortion present will cause the wave of plate current to deviate in form from the grid voltage. The following types of distortion may be present in any amplifier, either singly or together:

1. Amplitude distortion.
2. Frequency distortion.
3. Phase-shift distortion.

Amplitude distortion is due to the production of new frequencies in the output which were not present in the input. If the signal voltage e_s applied to the grid in Fig. 91 is sinusoidal, the plate current will be a distorted sine wave unless the portion of the characteristic covered by e_s is linear. Any distorted wave can be resolved into its fundamental and harmonics, and since the latter are frequencies which were not present in the input, new frequencies have been produced by the device itself. The signal voltage to be amplified is usually a complex wave consisting of a number of frequencies so that in addition to the harmonics of the impressed signal frequencies, the output will also contain the sums and differences of these various frequencies. As mentioned in Sec. 54 of the preceding chapter, the relation between the signal voltage and plate current can be represented by the series

$$I_p = I_0 + ae_s + be_s^2 + ce_s^3 + \cdots \tag{2}$$

For convenience we shall assume that the impressed signal is composed of only two frequencies and is given by

$$e_s = E_1 \sin \omega_1 t + E_2 \sin \omega_2 t \tag{3}$$

Substituting (3) in (2),

$$\begin{aligned} I_p &= I_0 + aE_1 \sin \omega_1 t + aE_2 \sin \omega_2 t + bE_1^2 \sin^2 \omega_1 t + bE_2^2 \sin^2 \omega_2 t \\ &\qquad + 2bE_1E_2 \sin \omega_1 t \sin \omega_2 t + \cdots \\ &= I_0 + \frac{bE_1^2}{2} + \frac{bE_2^2}{2} + aE_1 \sin \omega_1 t + aE_2 \sin \omega_2 t \end{aligned}$$

$$-\frac{bE_1^2}{2}\cos 2\omega_1 t - \frac{bE_2^2}{2}\cos 2\omega_2 t + bE_1E_2\cos(\omega_1 - \omega_2)t - bE_1E_2\cos(\omega_1 + \omega_2)t + \cdots \quad (4)$$

The sum and difference frequencies are more objectionable forms of distortion than the harmonics in the reproduction of music as the former are usually discordant. The higher powers of e_s in (2) give rise to additional frequencies.

Another form of amplitude distortion may be caused by the flow of grid current if the grid is allowed to become positive. When the source of e_s has appreciable internal impedance z_s, the positive half cycles will be reduced in amplitude by an amount

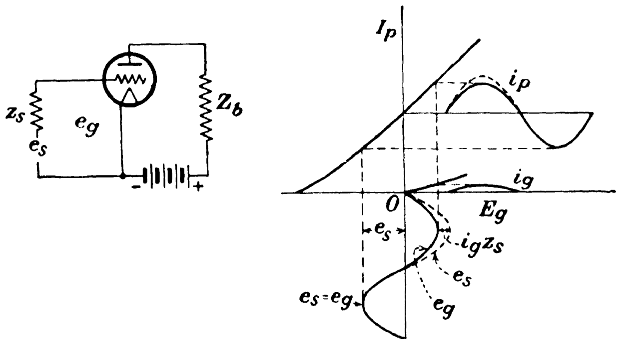


FIG. 92.—Amplitude distortion caused by the flow of grid current.

$i_g z_s$ which will distort the applied grid voltage in the manner shown in Fig. 92.

Frequency distortion is the unequal amplification of the various frequencies contained in the signal. It is due to the frequency characteristics of the input and output circuits associated with the tube. For example, if the plate-circuit impedance Z_b is chiefly inductive reactance, the value of this impedance at the lower frequencies will be smaller than at the higher values, which causes the voltage amplification to fall off at the lower frequencies.

Phase-shift distortion is the change in the relative phase relations of the different frequency components contained in the signal voltage. This shift in the phase of the constituent harmonics may occur without necessarily changing their amplitude relations, resulting in an output wave which is no longer identical in shape with that of the input wave. This is illustrated

in Fig. 93 for the theoretical case where the third harmonic is retarded in phase by an angle which is θ degrees more than the retardation imposed upon the fundamental. This type of dis-

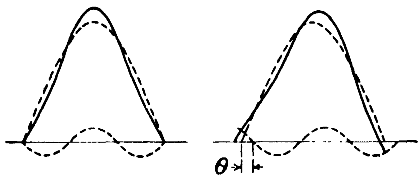


FIG. 93.—Effect of phase-shift distortion

tortion in amplifiers is also caused by the characteristics of the input and output circuits and is accompanied by frequency distortion as well.

Assume a tube having an amplification factor of μ and an internal resistance of r_p to have a pure inductance L_b in the plate

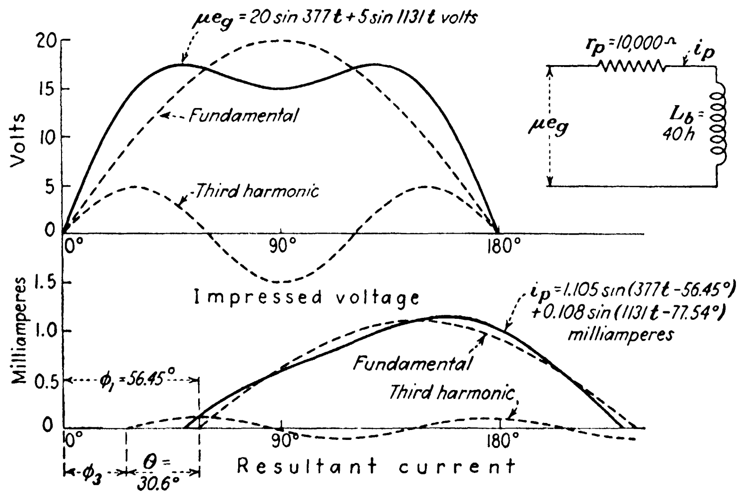


FIG. 94.—Phase-shift and frequency distortion produced in the plate current by an inductive load in the plate circuit of an amplifier tube.

circuit. The equivalent circuit is shown in Fig. 94, and the voltage acting in the plate circuit is assumed to be composed of a fundamental and a third harmonic given by the expression

$$\mu e_g = E_1 \sin \omega t + E_3 \sin 3\omega t$$

The fundamental component of the plate current i_{p1} will lag

behind the fundamental component of μe_g by the angle ϕ_1, which is given by

$$\phi_1 = \tan^{-1} \frac{\omega L_b}{r_p} \tag{5}$$

The third-harmonic component of the plate current i_{p3} will lag behind the third harmonic component μe_g by an angle

$$\phi_3 = \tan^{-1} \frac{3\omega L_b}{r_p} \tag{6}$$

Since 180 degrees for the third harmonic are only 60 degrees for the fundamental, the phase shift θ in fundamental degrees will be $\phi_1 - \frac{1}{3}\phi_3$.

Letting $\mu e_g = 20 \sin 377t + 5 \sin 1131t$ volts, $r_p = 10{,}000$ ohms, and $L_b = 40$ henrys, the fundamental component of plate current will be

$$i_{p1} = \frac{20 \sin (377t - \phi_1)}{\sqrt{r_p^2 + (\omega L_b)^2}} = 1.105 \sin (377t - 56.45°) \text{ milliamperes}$$

and

$$i_{p3} = \frac{5 \sin (1131t - \phi_3)}{\sqrt{r_p^2 + (3\omega L_b)^2}} = 0.108 \sin (1131t - 77.54°) \text{ milliamperes}$$

The phase shift θ will be

$$\theta = \phi_1 - \frac{1}{3}\phi_3 = 56.45° - \frac{77.54°}{3} = 30.6 \text{ fundamental degrees}$$

These values are shown to scale in Fig. 94. In addition to the distortion caused by phase shift, frequency distortion is also present, as the ratio of the amplitudes of the third harmonic and fundamental are no longer the same in the output as they were in the input.

Phase-shift distortion is not serious in audio amplifiers as the ear integrates the various components without regard to their phase relations as long as the shift in phase is not too great. In long telephone circuits used to transmit chain-broadcast programs, phase-shift distortion may be appreciable owing to the difference in the velocity of propagation of the various frequencies. This is compensated for by the use of suitable networks designed to have an inverse characteristic from that of the line.

The compensation results in a brief delay which may be observed by tuning a radio set to a frequency midway between two broadcasting stations which are transmitting the same chain program on adjacent channels. The difference in time delay between the two stations is clearly apparent on spoken announcements.

Phase-shift distortion would be highly objectionable in an amplifier designed to increase the sensitivity of an oscillograph because of the alterations that would be produced in the appearance of the wave. In fact, an oscillograph test is a convenient means of determining the amount of phase-shift distortion in an amplifier.

61. Amplifier with Resistance Load.—The effect of the load in the plate circuit is to straighten out the characteristic, as shown

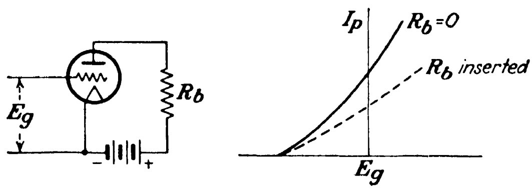


FIG. 95.—Effect of resistance in the plate circuit on the I_p-E_g characteristic.

in Fig. 95 for the case of a pure resistance in the plate circuit. As the grid voltage is increased from cut-off, the plate current rises and the voltage drop across R_b increases, causing a reduction in the plate voltage applied to the tube. Consequently, the new characteristic will lie below the curve for $R_b = 0$ and will approach the horizontal axis as R_b becomes enormously large. As this limiting position is a straight line, it is evident that the curvature in the characteristic must diminish as R_b is increased. Straightening the operating characteristic in this manner will greatly reduce the amplitude distortion. The dotted curve in Fig. 95 represents the actual working characteristic with the load R_b in the plate circuit and is spoken of as the *dynamic characteristic*. The characteristic of the tube itself with $R_b = 0$ is the *static characteristic*.

If an alternating voltage e_g is applied to the grid of the tube the alternating component of plate current will be

$$i_p = \frac{\mu e_g}{r_p + R_b} \tag{7}$$

The amplified output voltage is i_pR_b, so that the voltage amplification is

$$A_v = \frac{i_pR_b}{e_g} = \frac{\mu R_b}{r_p + R_b} \tag{8}$$

As R_b approaches infinity, A_v approaches μ of the tube as a limiting value. The increase in the voltage amplification as R_b is increased is given in the following table for a tube having $\mu = 10$

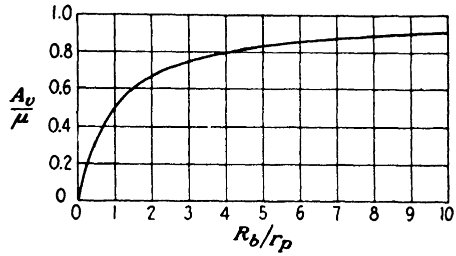


FIG. 96.—Variation of voltage amplification as a function of R_b/r_p.

and $r_p = 10{,}000$. Figure 96 shows the variation in amplification in terms of the ratio of R_b to r_p. The increase in amplification

R_b	A_v
1,000	0.909
10,000	5.00
100,000	9.09
1,000,000	9.90

is rather slow for values of R_b much above four or five times the value of r_p. Furthermore, large values of R_b will consume the greater portion of E_b in the form of I_bR_b drop, so that the plate voltage E_p across the tube may be only a small fraction of E_b. In order to secure high amplification with a resistance load in the plate circuit, it is necessary to use a tube having a high value of μ, or else use several stages of amplification in cascade.

The expression for the voltage amplification in (8) assumes that μ and r_p are constant as e_g varies, which is true if the amplitude of e_g is small. In many cases it is necessary to obtain as large an amplified output voltage as the tube and the allowable distortion will permit. In these cases e_g may be too large to permit this assumption to be made without appreciable error, so that a graphical solution is often used. This method of solution also has the added advantage of enabling the amount of

distortion to be determined, as will be shown later. The dynamic I_p-E_g characteristic shown as the dotted curve in Fig. 95 can be used, but this curve has to be constructed from the static characteristics of the tube for each particular value of load resistance that is tried. It is accordingly more convenient to use the static I_p-E_p characteristics, which will avoid this difficulty. The procedure is illustrated in Fig. 97 for the high-mu triode portion of a 2A6 tube, which has an amplification factor of 100.

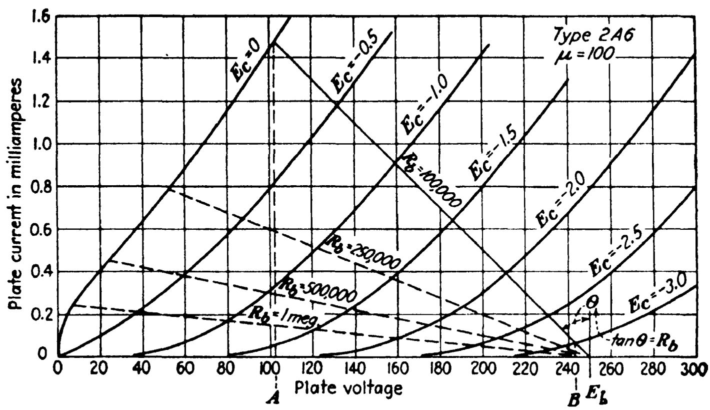


FIG. 97.—Graphical determination of the voltage amplification of a triode with a resistance load in the plate circuit.

The construction for a plate-supply voltage of $E_b = 250$ is as follows:

From E_b a line is drawn with a slope such that $\tan \theta = R_b$, taking into account the difference in current and voltage scales. The case illustrated is for a load of 100,000 ohms. Assuming a negative grid bias on the tube of $E_c = 1.5$ volts, the intersection of this load line with the characteristics $E_c = 0$ and $E_c = -3$ will give the maximum and minimum values of plate current. The grid potential will vary sinusoidally with time along the load line from zero to -3 volts. Owing to the voltage drop in R_b, the plate voltage across the tube will vary from a minimum value of OA to a maximum value of OB. The voltage across R_b is a maximum when the plate voltage is a minimum. The maximum voltage across R_b will be E_bA, while the minimum voltage will be E_bB. The useful or alternating component of this voltage is $\frac{1}{2}AB$, since the distance AB is equal to the double

amplitude of the output voltage. The voltage amplification is

$$A_v = \frac{\frac{1}{2}(B - A)}{e_g} = \frac{\frac{1}{2}(243 - 102)}{1.5} = 47.0$$

Using $\mu = 100$ and taking r_p as the cotangent of the slope of the characteristic $E_c = 1.5$ where it intersects the 100,000-ohm-resistance line, we find from the curve that $r_p = 102{,}500$ ohms. Substituting these values in (8),

$$A_v = \frac{100 \times 100{,}000}{102{,}500 + 100{,}000} = 49.4$$

which is about 5 per cent higher than the more accurate graphical determination. This inaccuracy is due chiefly to the assumption that r_p remains constant over the entire swing of grid voltage.

Load lines corresponding to other values of R_b are shown dotted in the diagram. In this way the performance of the amplifier for various loads may be readily determined. The amplitude distortion is diminished as the load resistance in the plate circuit is increased. The straightening effect of various values of load resistance upon the dynamic characteristic is shown in Fig. 98. From the amount of curvature still present in the characteristic for $R_b = 100{,}000$, it is evident that there will be appreciable amplitude distortion for a signal having an amplitude of 1.5 volts. To reduce this, either the signal must be limited to a smaller value, or else the value of R_b must be increased. Frequency and phase-shift distortion will be absent in this type of circuit and irregular direct-current impulses can be amplified, since the impedance in the plate circuit is assumed to be a pure resistance.

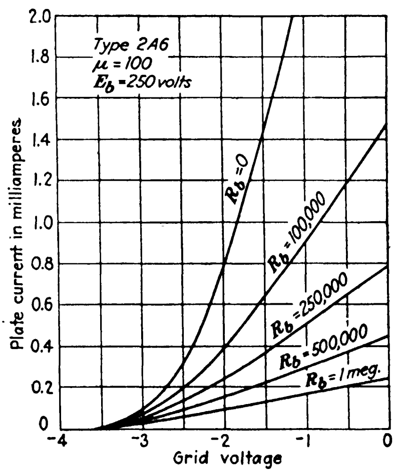


FIG. 98.—Dynamic characteristics of a triode showing the effect of resistance in the plate circuit upon the curvature of the characteristic.

62. Resistance-coupled Amplifiers.—The amount of gain afforded by a single stage of amplification similar to the circuit of Fig. 95 is not sufficient for many applications and several stages connected in cascade must be used. If the input circuit of the second stage were connected directly across the output resistance R_b of the first stage, the direct-current-voltage drop I_bR_b across this resistance would impose an excessive voltage on the grid of the second tube, which would bias it considerably beyond cut-off. This would require a very large C battery in the grid circuit of the second tube to offset this unwanted bias. If only alternating current is to be amplified, a blocking condenser

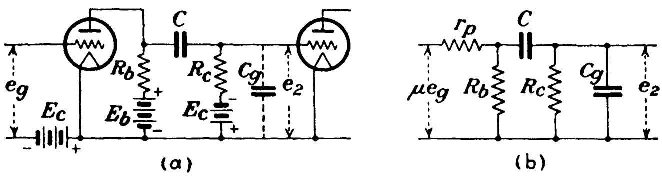


FIG. 99.—Resistance-coupled amplifier and its equivalent circuit.

in the grid lead of the second tube will serve to insulate the grid from this direct-current potential and at the same time offer a path of low impedance for the alternating signal voltage.

The circuit is shown in Fig. 99*a*. The blocking condenser C would insulate the grid of the second tube and allow it to accumulate a charge, so that a "grid leak" R_c is provided which enables the proper C bias to be imposed on the tube. The equivalent circuit is shown in Fig. 99*b*. The condenser C_g is the input capacitance of the second tube, augmented by the stray capacitance of the wiring, etc. The value of C_g is by no means negligible and depends upon the load in the plate circuit of the second tube, and also upon the interelectrode capacitances of this tube. It will later be shown that this input capacitance is approximately constant and independent of the frequency for values below 150 kc in the case of a tube having a pure resistance in its plate circuit. The expression for this capacitance is

$$C_g = C_{gf} + C_{gp}\left(1 + \frac{\mu R_b}{r_p + R_b}\right) \tag{9}$$

where C_{gf} = grid-filament capacitance.

C_{gp} = grid-plate capacitance.

All the quantities in (9) pertain to the second tube. The last

term will be recognized as the expression for the voltage amplification A_v of the second stage. If the second tube is the same as discussed in the example in Sec. 61, for which $C_{gf} = C_{gp} = 1.7$ $\mu\mu$f, the input capacitance will be

$$C_g = 1.7 + 1.7(1 + 49.4) = 87.4 \ \mu\mu\text{f}$$

There will also be some resistance associated with C_g owing to the resistance in the plate circuit of the tube and also to the dielectric loss associated with the various interelectrode capacitances, which will be discussed in Sec. 90, Chap. VIII. This resistance can usually be neglected at audio frequencies.

At low values of frequency the effect of shunting C_g across R_c is negligible, but as the frequency increases the effect becomes more pronounced until at very high frequencies R_c is virtually short-circuited. This causes a reduction in the amplification for the higher values of audio frequencies. With C_g a fixed amount, the lower the value of R_c, the higher the frequency at which the drop in amplification becomes pronounced. However, R_c is in parallel with R_b in effect, as the reactance of the blocking condenser C can be neglected at all but the lower frequencies, so that as R_c is made smaller the voltage amplification will be reduced. For moderate values of frequency the alternating-current load impedance may be regarded as being composed of R_b and R_c in parallel. The reactance of C_g is usually too high in this range of frequencies to be of consequence.

At very low frequencies the reactance of the blocking condenser becomes of importance, since the output of the first tube is impressed across C and R_c in series. As $1/\omega C$ becomes comparable to R_c, less of the output voltage across R_b is applied to the grid of the second tube and the amplification again falls off, approaching zero as the frequency is reduced. The insulation resistance of the blocking condenser must be extremely high as E_b is impressed across it in series with R_b and R_c. If there is appreciable leakage in the condenser, a current will flow through R_c which may be sufficient to place a positive bias on the grid of the second tube, especially if R_c is comparatively large. For this reason a mica condenser is usually preferred. Its capacitance is usually in the vicinity of 0.1 μf.

The values of R_b and R_c depend upon the type of tube used and the range of frequencies over which constant amplification is

desired. The grid leak R_c will have a value of from one to ten times R_b, the lower values being more usual.

The calculation of the voltage amplification from the equivalent circuit of Fig. 99*b* for the entire audio-frequency range is

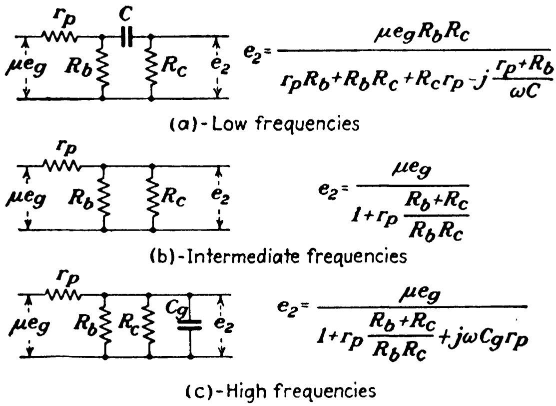


FIG. 100.—Approximate equivalent circuits of a resistance-coupled amplifier.

rather tedious, so that the approximate equivalent circuits of Fig. 100 may be used for such calculations with sufficient accuracy for most purposes. The voltage amplification per stage of a typi-

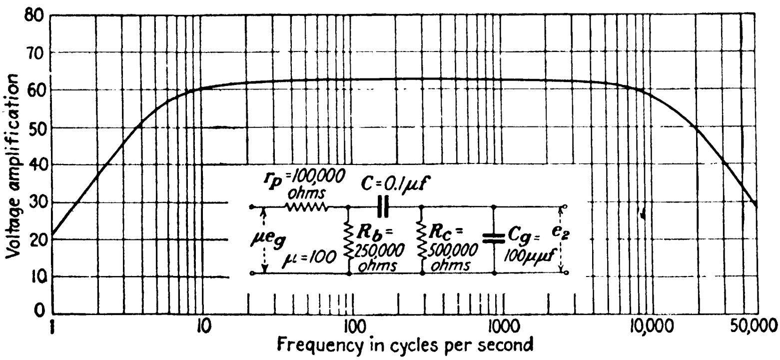


FIG. 101.—Variation of voltage amplification with frequency in a typical resistance-coupled amplifier.

cal resistance-coupled amplifier is shown in Fig. 101 for a 2A6 tube with R_b and R_c 0.25 and 0.5 megohm, respectively. The values of the circuit constants are given in the diagram.

The voltage amplification may be determined graphically by means of the construction shown in Fig. 102. This differs from

the construction used in Fig. 97 in that the plate-circuit resistance for the alternating component of plate current i_p is lower than for the direct-current component I_b. The direct-current resistance is R_b while the alternating-current resistance, neglecting the effects of C and C_g, is $\frac{R_b R_c}{R_b + R_c}$. From the point E_b corresponding to the value of plate-supply voltage used, a line whose slope is $\tan^{-1} R_b$ is drawn until it intersects the characteristic correspond-

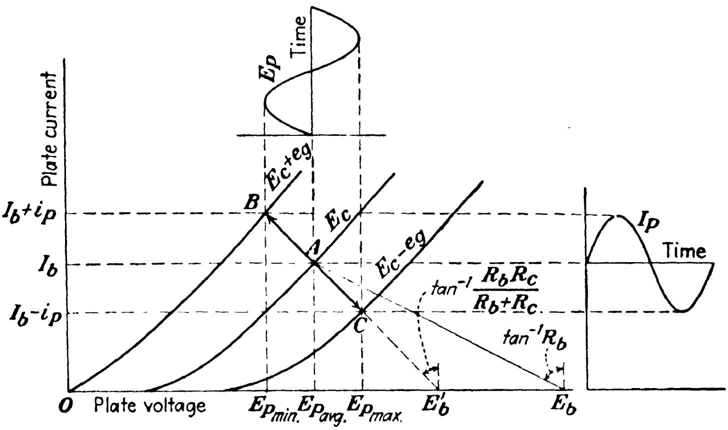


FIG. 102.—Graphical construction for the determination of the voltage amplification of a resistance-coupled amplifier when the alternating-current resistance of the plate circuit is lower than the direct-current resistance.

ing to the value of C bias used, at point A. Through point A another line is drawn whose slope is equal to $\tan^{-1} \frac{R_b R_c}{R_b + R_c}$. The distance BC represents the double amplitude of the signal voltage, e_g. This line intersects the plate voltage axis at E_b', which would be the value of plate-supply voltage needed if the direct-current plate-circuit resistance were equal to the alternating-current value. The voltage amplification will be

$$A_v = \frac{E_{p\max} - E_{p\min}}{2e_g} \tag{10}$$

63. Impedance-coupled Amplifiers.—If an impedance, such as a choke coil, is connected in the plate circuit of a triode, as shown in Fig. 103, the alternating component of the plate current is

$$i_p = \frac{\mu e_g}{r_p + Z_b} = \frac{\mu e_g}{r_p + R_b + j\omega L_b} \tag{11}$$

The voltage e_2, in absolute magnitude, is given by

$$e_2 = i_p Z_b = \frac{\mu e_g \sqrt{R_b^2 + \omega^2 L_b^2}}{\sqrt{(r_p + R_b)^2 + \omega^2 L_b^2}} \tag{12}$$

and the voltage amplification is

$$A_v = \frac{e_2}{e_g} = \frac{\mu \sqrt{R_b^2 + \omega^2 L_b^2}}{\sqrt{(r_p + R_b)^2 + \omega^2 L_b^2}} \tag{13}$$

If the resistance of the coil is negligible compared to the reactance, as is usually the case, (13) becomes

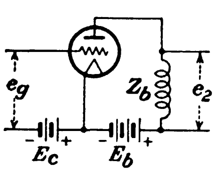


FIG. 103.—Amplifier with an inductive load.

$$A_v = \frac{\mu \omega L_b}{\sqrt{r_p^2 + \omega^2 L_b^2}} \tag{14}$$

As the frequency increases, ωL_b becomes large compared with r_p and the amplification approaches μ of the tube as a limiting value. This is similar to the case of the resistance-coupled amplifier when R_b was made very large compared with r_p. The increase in voltage amplification as ωL_b is increased is given in the following table for a tube having $\mu = 10$ and $r_p = 10{,}000$. It

ωL_b	A_v
1,000	0.995
10,000	7.07
100,000	9.99

will be observed by comparing these values with those on page 151 for the case of a resistance in the plate circuit that the amplification per ohm of plate load is greater in the case of the inductive load. This is due to the fact that the denominator of (14) is the vector sum of the plate resistance and the plate load instead of the scalar sum.

The above expressions for the voltage amplification are functions of the frequency and it might seem that this type of amplifier would be objectionable from the standpoint of frequency distortion. This is not necessarily the case, since the amplification is substantially constant and equal to μ if ωL_b is very large compared to r_p. Consequently, if we choose a value of L_b so that ωL_b is large compared with r_p for the lowest frequency with which we are concerned, the amplification will be practically constant for all higher frequencies, provided that the capacitance

shunted across L_b is negligible. This is illustrated in Fig. 104. With a value of L_b of 100 henrys the amplification is practically constant for frequencies above 60 cycles.

In order to obtain high amplification per stage, it is necessary to use a tube having a high value of μ. These tubes have a correspondingly high plate resistance so that rather large values of L_b are necessary to extend the range of constant amplification into the lower frequencies. The large number of turns on the

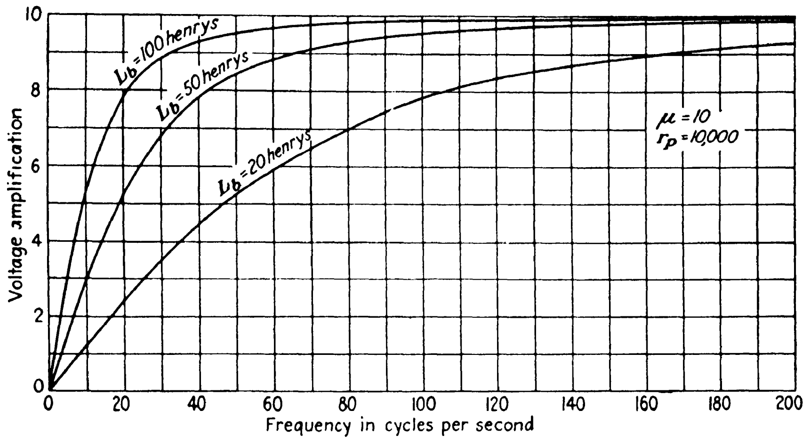


FIG. 104.—Variation of voltage amplification at low frequencies for various values of inductance in the plate circuit.

coil needed to secure the high inductance required gives rise to saturation difficulties in the iron core, owing to the presence of I_b. Furthermore, the distributed capacitance is apt to be high, which augments the capacitance C_g shunted across the coil in the case of a cascade amplifier. The tube-input capacitance C_g will be of the same order of magnitude in a reactance-coupled amplifier as it was in the resistance-coupled case. The total capacitance shunted across L_b will constitute a parallel-resonant circuit of very high impedance at some particular frequency. This merely insures a closer approach to the maximum theoretical value of amplification, but at frequencies above resonance the impedance begins to fall and the amplification curve begins to droop in a fashion similar to the resistance-coupled amplifier. At resonance the impedance of the parallel circuit is usually so high that the resistance of the grid leak R_c becomes the limiting factor. The amplification can usually be determined for this resonant fre-

quency by the use of (8) for the resistance-coupled case, by substituting R_c for R_b. The circuit of a typical stage of an impedance-coupled amplifier is shown in Fig. 105.

Impedance-coupled amplifiers have one advantage over resistance coupling in that the resistance of the coil to direct current is relatively small, so that the average plate voltage impressed

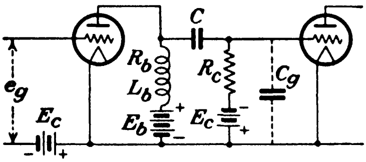


FIG. 105.—Impedance-coupled amplifier.

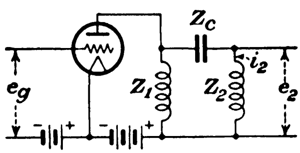


FIG. 106.—Double impedance-coupled amplifier.

on the tube is almost equal to E_b. This allows lower values of plate-supply voltage to be used. The apparent resistance of the coil to alternating current is much greater than the direct-current resistance because of the core loss.

In place of the grid leak R_c in Fig. 105, a coil Z_2 of high inductance may be used, as in Fig. 106. At some value of frequency Z_2 and Z_c will constitute a series-resonant circuit shunted across Z_1, resulting in a resonant rise in the voltage e_2 across the coil Z_2. The vector expression for this voltage, neglecting any other impedance, such as tube-input capacitance that may be shunted across Z_2, is

$$e_2 = i_2 Z_2 = \frac{\mu e_g Z_1}{r_p Z_1 + Z_1(Z_2 + Z_c) + r_p(Z_2 + Z_c)} Z_2$$

$$= \frac{\mu e_g Z_2}{r_p\left(1 + \frac{Z_2 + Z_c}{Z_1}\right) + Z_2 + Z_c} \qquad (15)$$

At resonance, $Z_2 + Z_c = R_2$, which is the resistance of the coil Z_2, and (15) becomes

$$e_2 = \frac{\mu e_g Z_2}{r_p\left(1 + \frac{R_2}{Z_1}\right) + R_2} \qquad (16)$$

If Z_1 is very large compared with R_2, the approximate value of e_2 will be

$$e_2 = \frac{\mu e_g Z_2}{r_p + R_2} \qquad (17)$$

By making Z_2 greater in magnitude than $r_p + R_2$, it is possible to make the voltage amplification exceed μ of the tube, which has been heretofore the maximum value obtainable. The amplification curve of the ordinary type of reactance-coupled amplifier tends to fall off at the lower frequencies, as shown in Fig. 104. By choosing Z_c and Z_2 so that they are in resonance at some low value of frequency, the drop in amplification which normally occurs in this region may be made to have a resonant rise superimposed upon it.

64. Amplifier with Inductive Load. The circuit of Fig. 106 is commonly used to couple a load, such as a loudspeaker, to an amplifier when it is necessary to exclude the direct-current component of plate current I_b from the speaker windings. This form of connection is often called shunt or parallel feed, as the plate-supply voltage is connected across the tube through a suitable choke coil instead of in series with the load and the tube. The connections are shown in Fig. 107. The blocking condenser prevents the battery current from flowing through the load Z_b, while the high inductance L of the choke coil practically excludes the alternating component of plate current and confines it to the load. The value of inductance required does not need to be so large as was the case with the impedance-coupled voltage amplifier since these applications are mostly concerned with power output and use tubes having a much lower value of r_p. The resistance R of the choke coil is very small compared to the internal resistance of the tube and the average value of plate voltage supplied to the tube is approximately E_b. The blocking condenser is usually made large enough so that its reactance is small in comparison to Z_b. With inductive loads the reactance of C serves to annul partially some of the inductive reactance.

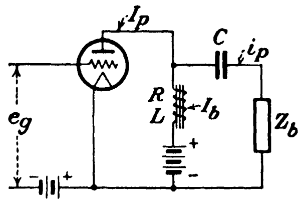


FIG. 107.—Amplifier using parallel feed.

The voltage and current relations for this type of circuit are shown in Fig. 108, for the case where Z_b is a pure resistance R_b and the reactance of C is negligibly small. The construction differs from that of Fig. 102 in that the resultant resistance in the plate circuit is now higher for alternating current than for direct current. Note that the maximum value of plate voltage

is greater than E_b. This is caused by the e.m.f. of self-induction $L\frac{di}{dt}$ in the choke coil, which alternately is added to, and subtracted from, E_b.

When the plate circuit impedance contains reactance, the path followed by the grid voltage is an ellipse as shown in Fig. 109. The path taken around the ellipse will be clockwise, as shown by the arrows, if X_b is inductive. If X_b had been capacitive, a counterclockwise direction would have resulted, with corresponding phase shifts of I_p and E_p. The major axis of the ellipse makes an angle with the vertical whose tangent is $|Z_b|$. The center of the ellipse is located on a line drawn from E_b making an angle

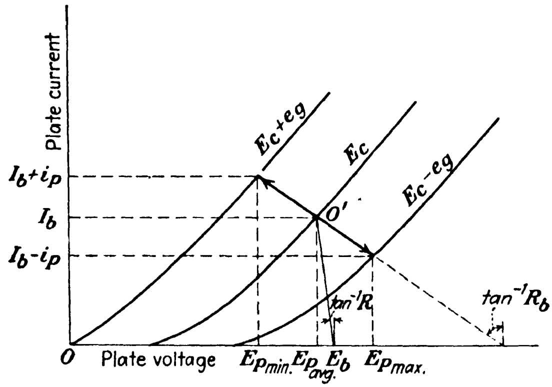


FIG. 108.—Current and voltage relations for the circuit of Fig. 107 when the load Z_b is a pure resistance, R_b.

with the vertical whose tangent is R_b. The example assumes that the alternating-current resistance of the plate load is the same as its direct-current resistance. This line intersects the elliptical path at its points of tangency with the horizontal lines representing the maximum and minimum values of plate current. If the direct-current resistance of the load had been lower than the alternating-current value, as is often the case, a third line through the center could be drawn making an angle with the vertical whose tangent is the direct-current resistance. The intersection of this last line with the plate-voltage axis would then be equal to E_b and the points of intersection of the other lines with this axis would represent fictitious values. The ellipse degenerates into a straight line when the reactance of the load is made zero. The path of operation becomes a circle in

the case of a pure reactance load equal to r_p in magnitude. A perfect ellipse or a circle is obtained only when the working portions of the static characteristic are linear and parallel.

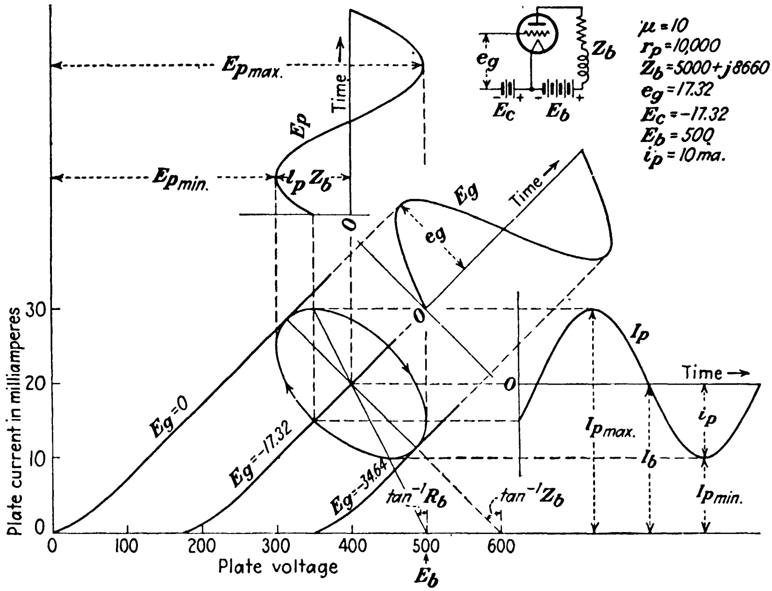


FIG. 109.—Path of operation on the I_p-E_p characteristic for an inductive plate load.

In most cases the reactance of the load is small compared to its resistance, so that the path of operation may be assumed to be a straight line instead of an ellipse with sufficient accuracy for most purposes.

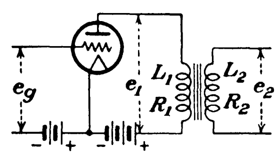


FIG. 110.—Transformer-coupled amplifier.

65. Transformer-coupled Amplifiers. In the preceding cases the amplification per stage has been limited to a value approaching μ of the tube, with the single exception of the double impedance-coupled amplifier. In order to increase the gain, it is possible to use a suitable step-up transformer, as shown in Fig. 110. Assuming the secondary to be on open circuit and perfect coupling to exist between the two windings, the secondary voltage e_2 will be the ratio of transformation times the voltage e_1 across the primary. The primary voltage under these assumptions will be the same as for the

impedance-coupled case given by (12), and the seconaary voltage will be

$$e_2 = \frac{\mu e_g \sqrt{R_1^2 + \omega^2 L_1^2}}{\sqrt{(r_p + R_1)^2 + \omega^2 L_1^2}} a \tag{18}$$

where a is the ratio of transformation and is equal to N_2/N_1. The inductance of the primary must be large if frequency distortion in the form of loss of amplification at the low frequencies is to be avoided. If R_1 can be neglected in comparison with ωL_1 and with r_p, the voltage amplification is

$$A_v = \frac{\mu a \omega L_1}{\sqrt{r_p^2 + \omega^2 L_1^2}} \tag{19}$$

The amplification curve for low frequencies will be similar in appearance to Fig. 104 if the ordinates of the curve are multiplied by the ratio of transformation a. At frequencies such that ωL_1 is very large compared to r_p of the tube used,

$$A_v = \mu a \tag{20}$$

The above approximations are valid only for frequencies below about 500 cycles. For higher frequencies the effects of distributed capacitance, tube-input capacitance, and primary and secondary leakage reactance, become of increasing importance. The equivalent circuit of a transformer may be regarded as an ideal transformer which has the various imperfections of the actual transformer inserted in its external primary and secondary circuits.

An ideal transformer is one whose primary and secondary impedances are infinite inductive reactances; hence there will be no losses or magnetizing current. The coefficient of coupling is unity, and therefore $M = \sqrt{L_p L_s}$, and is also infinite. The ratio of transformation will be $a = N_s/N_p = \sqrt{L_s/L_p}$, since the inductance varies as the square of the number of turns. It should be remembered that an infinite term is one which is larger than the largest assignable quantity and that infinite quantities are not necessarily equal. The currents in the primary and secondary of an ideal transformer can be obtained from (14) and (15) of Sec. 34 for an actual transformer, by inserting the above relations and neglecting finite terms in comparison with infinite terms. These equations for an ideal transformer are

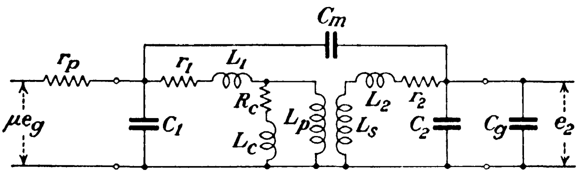
$$I_p = \frac{EZ_s}{Z_pZ_2 + Z_sZ_1} = \frac{E}{\frac{Z_p}{Z_s}Z_2 + Z_1} = \frac{E}{\frac{1}{a^2}Z_2 + Z_1} \tag{21}$$

and in a similar fashion

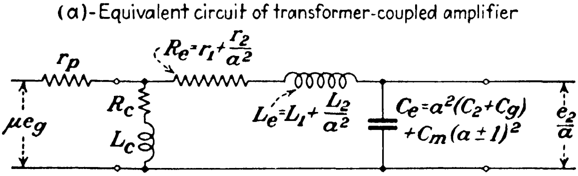
$$I_s = \frac{E}{\frac{1}{a}Z_2 + aZ_1} \tag{22}$$

From (21) the impedance of the secondary load Z_2 as viewed from the primary terminals of the transformer will be Z_2/a^2. In other words, an ideal transformer will step down or step up the secondary load impedance by an amount equal to $1/a^2$ depending on whether the ratio of transformation a is greater or less than unity.

The equivalent circuit of a transformer-coupled amplifier is shown in Fig. 111a, where L_p and L_s represent an ideal transformer



(a)-Equivalent circuit of transformer-coupled amplifier



(b)-Approximate equivalent circuit

FIG. 111.—Equivalent circuits of transformer-coupled amplifier.

as defined above. The primary-winding resistance, leakage inductance, and distributed capacitance are r_1, L_1, and C_1. The similar secondary quantities are represented by r_2, L_2, and C_2. The core loss and magnetizing components of the no-load current flow through R_c and L_c. The condenser C_m represents the capacitance between the primary and secondary windings, while C_g is the input capacitance of the second tube.

Reducing the secondary quantities to their equivalent primary values results in the circuit of Fig. 111b, with the added approximation that C_m is equivalent to a shunt capacitance whose value

is $C_m(a \pm 1)^2$. The choice of sign depends on whether the capacitive coupling is aiding or opposing the magnetic coupling. At low frequencies, R_e, L_e, and C_e may be neglected, but at higher frequencies these approach series resonance, causing a resonant rise in the output voltage across C_e. The height of this resonant peak is governed by the value of R_e, and the frequency at which it occurs, by L_e and C_e. At frequencies in the vicinity of resonance R_c and L_c may be neglected, as the impedance of this circuit is very high compared to the impedance of R_e, L_e, and C_e in series. The primary distributed capacitance C_1 should be inserted in place of R_c and L_c for frequencies much above the resonant frequency of L_e and C_e.

The voltage applied to the grid of the second tube for frequencies above the value for which (18) is accurate, from Fig. 111*b*, will be given by the vector equation

$$e_2 = \frac{\mu e_g a}{j\omega C_e\left(r_p + Z_e + \frac{r_p Z_e}{Z_c}\right)} \tag{23}$$

where

$$Z_c = R_c + j\omega L_c$$

$$Z_e = R_e + j\left(\omega L_e - \frac{1}{\omega C_e}\right).$$

For the higher frequencies where Z_c is very large, the last term in the denominator of (23) may be neglected. For very high frequencies Z_c in the last term should be replaced by the reactance of the condenser C_1.

66. Transformer Characteristics.—Figure 112 shows the variation of amplification with frequency for several different types of transformers. Curve 1 is for a small type widely used in receiving sets in 1926. Insufficient primary inductance (about 15 henrys) is responsible for the droop in the curve at the lower frequencies. Curves 2 and 3 are for larger transformers typical of the designs a few years later. The primary inductance has been increased giving better low-frequency response. The high resonant peak of transformer 2 is due chiefly to too low a value of secondary resistance. This particular design used No. 36 enameled wire on both primary and secondary instead of the usual No. 40. The height of the resonant peak can be materially

reduced by increasing the secondary resistance, so that wire sizes smaller than No. 40 are often used for the secondary. Production difficulties are greater with the smaller sizes of wire. A high resistance of 0.5 to 1 megohm shunted across the secondary is effective in removing the resonant peak from the characteristic.

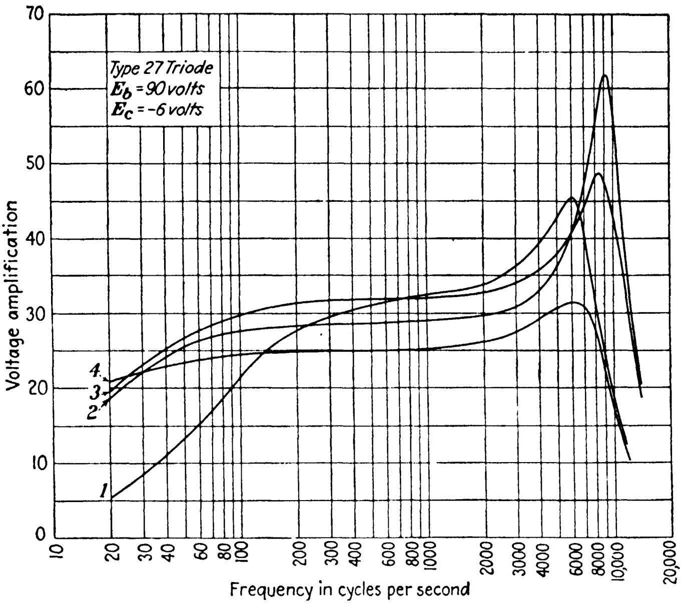


FIG. 112.—Voltage amplification per stage of a transformer-coupled amplifier for various types of transformers.

The entire curve is lowered but the amplification is more nearly constant over the frequency range.

Curve 4 is representative of a modern design of good quality. High primary inductance is secured by the use of an alloy core. The value in this case is 116 henrys with a direct current of 2.6 ma in the primary. As will be recalled, the alternating-current inductance of a coil with a magnetic core is reduced as the direct-current excitation is increased, as was shown in Chap. III, Fig. 53. The top curve in this figure was obtained with the same transformer that was used for Curve 4 in Fig. 112. It is, therefore, desirable to keep the value of I_b as small as possible by the use of a moderate value of plate voltage on the tube and a sufficient value of negative bias on the grid. This is feasible if the

magnitude of the alternating voltage that is to be amplified is small. If the alternating voltage is relatively large, insufficient plate voltage will give rise to appreciable amplitude distortion. The incremental permeability of the iron increases somewhat with the amplitude of the alternating current, which in part helps to offset the reduced inductance caused by the increase in I_b when larger plate voltages are used. The use of a suitable air gap in the magnetic circuit will reduce the effect of the direct-current magnetization, as pointed out in Chap. III.

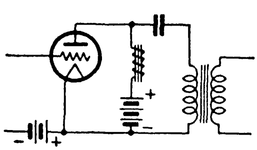


FIG. 113.—Transformer-coupled amplifier using parallel feed.

The undesirable effect of direct current in the transformer primary may be avoided by the circuit of Fig. 113. This circuit is similar to the double impedance-coupled amplifier of Fig. 106 and there will be a similar resonant rise in amplification at some low value of frequency. The choke coil in series with the plate-supply voltage must also have a very high inductance, so that the difficulties present in the transformer design are now transferred to the choke coil. However, all of the winding space can be utilized on the choke coil, whereas the primary winding can only occupy a portion of the winding space in the case of the transformer.

In power transformers the space occupied by the primary and secondary windings is about the same, since the winding having the greater number of turns will use a proportionately smaller size of wire. In audio-frequency transformers the size of the wire used on the secondary is already as small as practical considerations will permit, and if it is desired to increase the ratio of transformation, assuming the winding space and core size to be fixed, the only way in which this can be accomplished is to reduce the number of primary turns as the secondary turns are increased. This will reduce the primary inductance and impair the performance at low frequencies. In addition, the coefficient of coupling is apt to decrease with a greater number of secondary turns, so that ratios of transformation above 4:1 or 5:1 are difficult to secure without some sacrifice in performance. A 3:1 ratio represents the usual practice.

The secondary distributed capacitance, together with the input capacitance of the second tube, resonates with the equivalent

leakage inductance of the transformer to produce the characteristic peak in the amplification curve at high frequencies. At frequencies above this value the amplification falls rapidly. In order to extend the useful range of amplification the distributed capacitance and leakage inductance must both be kept small. Low leakage inductance means a high coefficient of coupling between primary and secondary, which can be obtained by more closely interleaving the two windings, as is done in power transformers to secure better regulation. However, this increases the capacitance C_m between the two windings, which is not

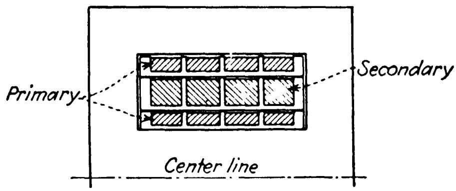


FIG. 114.—Audio-frequency transformer with sectionalized winding to reduce the distributed capacitance.

altogether desirable. The coupling coefficient in audio-frequency transformers will ordinarily range from 0.99 to perhaps 0.999, but with a primary inductance of 50 to 100 henrys and a secondary inductance which will be a^2 greater, the value of L_e referred to the primary will be in the vicinity of a henry, in spite of the high coupling coefficient. The value of C_2 will be usually from 50 to 100 $\mu\mu$f, and C_1 will be roughly a times greater, so that the resonant peak will ordinarily occur somewhere between 5000 and 10,000 cycles. The use of alloy cores which enable high inductance to be obtained with fewer turns helps to overcome these difficulties.

The distributed capacitance can be reduced to some extent by winding the primary and secondary in sections on molded spools which fit over one another, as illustrated in Fig. 114. The primary is wound in two parts, one inside and one outside the secondary, which reduces the leakage reactance.

67. Comparison of Amplifiers.—Transformer-coupled amplifiers are the most widely used type and can be designed to give substantially uniform amplification over the ordinary range of audio frequencies. They are capable of producing higher amplification per stage than the impedance-coupled type, unless tubes

having a high amplification factor are used in the latter. In this case very high values of coupling inductance must be used. The resistance-coupled amplifier is capable of delivering constant amplification over a wider frequency range than the two other types, particularly in the lower frequencies. It is cheaper to construct and is lighter and more compact than the other types. With the recent advent of various types of tubes having very high amplification factors the former limitation of low gain per stage is no longer present. However, these tubes are apt to be considerably more microphonic than those having a lower value of μ, so that mechanical vibration must be avoided if the output is to be free from noise. Microphonic noise is caused by the vibration of the tube elements which causes small variations to take place in the distances between them. This affects the amplification factor and is equivalent to impressing a very small alternating voltage on the grid of the tube.

When several stages of amplification are used in cascade, the total amplification will be the product of the amplifications of the individual stages. If all the stages are identical, the total amplification will be

$$A_t = A^n \tag{24}$$

where n is the number of stages. With like stages any frequency distortion present in the individual characteristic will be magnified in the overall characteristic, so that the amplification characteristic of each individual stage must be much flatter than the required overall characteristic.

68. Amplification Expressed in Decibels.—The voltage-amplification curves of Fig. 112 show considerable differences in gain between the various transformers at certain frequencies. The question arises as to what extent would these differences be apparent to the ear if a comparative audition test were made at some particular frequency? The sensation of loudness is rather difficult to measure in terms of physical values. Numerous tests indicate that the sensation of loudness is proportional to the logarithm of the stimulus, so that a logarithmic scale of ordinates in Fig. 112 would more nearly approximate the performance of the amplifier as interpreted by the ear.

In telephone communication a unit known as the *decibel* (db) is used to express the gain or loss in power due to the insertion of

a particular piece of apparatus in a circuit, and is equal to 10 times the common logarithm of the ratio of the two powers, or

$$\text{Number of decibels} = 10 \log_{10} \frac{P_2}{P_1} \tag{25}$$

where P_1 is the power supplied to the circuit before the insertion of the device and P_2 the power supplied after the device has been inserted. Since the power output to the circuit is proportional to the square of the voltage impressed across it, (25) can also be written

$$\text{Number of decibels} = 20 \log_{10} \frac{E_2}{E_1} \tag{26}$$

The gain produced by an amplifier can evidently be expressed in decibels as

$$A_{db} = 20 \log_{10} \frac{e_2}{e_g} = 20 \log_{10} A_v \tag{27}$$

The size of this unit is convenient in that a change in sound energy of 1 db is about the minimum amount that can be detected by the average ear.

The decibel is also used to express energy levels in terms of an arbitrary standard of reference or "zero level." This reference level in communication measurements is now taken as 0.001 watt, although other values have been used. In rating microphones, zero level is taken as an output of 1 volt (on open circuit) for a sound wave having a pressure of 1 dyne per square centimeter (1 bar). Thus, if a particular microphone is rated as −36 db, it will require an amplifier having a gain of +36 db to raise the output to zero level. This may be accomplished by a two-stage amplifier having an amplification of 18 db per stage, or else a three-stage amplifier having an amplification of 12 db per stage. Since the decibel is a logarithmic unit, the total gain of an amplifier can be conveniently found by adding together the gains of the individual stages. Computations relative to amplifiers and attenuating networks are very much simplified by the use of this type of unit.

69. Measurement of Voltage Amplification.—The voltage amplification of an amplifier can be readily measured by the arrangement of Fig. 115. The input voltage is determined by sending a known current from an audio-frequency oscillator

through a suitable voltage divider or attenuator. If the amplifier under test is provided with an input transformer, the resistance R_1 must be kept small in comparison with the input impedance, or the input voltage will no longer be IR_1. The output may be read directly by a vacuum-tube voltmeter. The amplifier should be terminated in a load impedance Z equal to the normal load on the amplifier. This is usually a pure resistance. In case the output of the amplifier is intended to be impressed on the

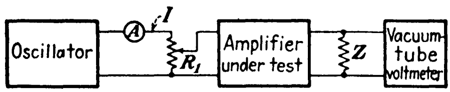


FIG. 115.—Circuit for the determination of voltage amplification.

input terminals of another vacuum tube, as in the case of a test on an individual stage, Z plus the input capacitance of the vacuum-tube voltmeter should be equal to the input capacitance of the following tube. The input capacitance of the voltmeter is usually much lower than the following amplifying tube because of the differences in the plate loads of the two tubes. The correct value of capacitance is of consequence only at the higher audio frequencies. At the lower frequencies it can be disregarded.

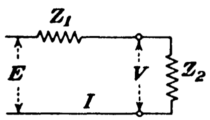


FIG. 116.—Source of e.m.f. E of internal impedance Z_1 with a load impedance Z_2.

70. Considerations for Maximum Power. In any amplifier we may be concerned with securing either the maximum voltage, power, or current. The first two conditions are the more usual, although the last requirement is occasionally met with in special applications.

Before considering amplifiers for these various requirements, let us investigate the general case of a source of e.m.f. E having an internal impedance Z_1 and a load connected across this source which has an impedance Z_2, as shown in Fig. 116. The current is

$$I = \frac{E}{Z_1 + Z_2} \tag{28}$$

If the current is to be made a maximum, assuming the load to be the variable, Z_2 should be made as small as possible.

The voltage across the load is

$$V = IZ_2 = \frac{EZ_2}{Z_1 + Z_2} \tag{29}$$

Considering Z_2 again to be the variable, V will be a maximum and equal to E when Z_2 is made infinite. Therefore for maximum voltage output the impedance of the load should be made as large as possible. This requirement applies to the voltage amplifiers just considered.

If the power output is to be a maximum, the conditions will depend upon what element of the circuit is variable. Assuming Z_1 and Z_2 to be pure resistances the power output will be

$$P = \frac{E^2 R_2}{(R_1 + R_2)^2} \tag{30}$$

If the resistance of the load is the independent variable, the power will be a maximum when $dP/dR_2 = 0$. Differentiating (30) with respect to R_2 and equating the result equal to zero gives us

$$R_1 = R_2 \tag{31}$$

as the necessary condition, which means that the resistance of the load must be made equal to the internal resistance of the source. But if the internal resistance of the source had been variable, with the load resistance fixed in value, maximum power would have been obtained when R_1 was made as small as possible.

Substituting (31) in (30), the expression for the maximum power when the load is the variable is

$$P_{\text{max}} = \frac{E^2}{4R_1} \tag{32}$$

This expression is useful in determining the maximum possible power a given source of e.m.f. is capable of delivering. Thus, a storage battery having an e.m.f. of 6 volts and capable of delivering 1200 amp. on short circuit—from which we find the internal resistance to be 0.005 ohm—from (32), is able to deliver a maximum output of 1800 watts.

If both Z_1 and Z_2 possess reactance as well as resistance, the power absorbed by the load will be

$$P = I^2 R_2 = \frac{E^2 R_2}{(R_1 + R_2)^2 + (X_1 + X_2)^2} \tag{33}$$

The conditions for maximum power output will depend upon which term is made the independent variable. If only R_2 of the

load is varied and X_2 remains fixed, we find by differentiation that

$$R_2 = \sqrt{R_1^2 + (X_1 + X_2)^2} \tag{34}$$

In other words, the resistance of the load must be made equal to the magnitude of the impedance of the remainder of the circuit. Since X_2 remains fixed, it is viewed by R_2 as being a portion of the internal impedance of the source.

Making X_2 the variable and holding R_2 constant, (33) is found to be a maximum when

$$X_2 = -X_1 \tag{35}$$

which means that if X_1 is inductive, X_2 should be made capacitive so that the total reactance in the circuit is zero. If this is not feasible, then X_2 should be made as small as possible.

In most cases both R_2 and X_2 will vary as the impedance of the load is changed. Suppose the load is some form of coil whose winding space is fixed, as the moving coil in a dynamic-type of loud-speaker. In order to increase the impedance of the coil it is necessary to use a greater number of turns of smaller wire. For example, if the number of turns were to be doubled, wire of half the present cross-sectional area would have to be used. The resistance would then be four times as great, and since the inductance varies as the square of the number of turns, the reactance of the coil would also be four times its former value.

Loads of this kind will therefore have a fixed ratio of X_2 to R_2, so that Z_2 can be expressed by

$$Z_2 = R_2 + jQR_2 \tag{36}$$

where $Q = X_2/R_2$.

The expression for the power in the load in this case is

$$P = \frac{E^2R_2}{(R_1 + R_2)^2 + (X_1 + QR_2)^2} \tag{37}$$

Differentiating this expression with respect to R_2, the power absorbed by the load is found to be a maximum when

$$\sqrt{R_1^2 + X_1^2} = R_2\sqrt{1 + Q^2} \qquad \text{or} \qquad |Z_1| = |Z_2| \tag{38}$$

In this case the impedance of the load should be made equal in magnitude to the impedance of the source.

If the impedance of the source is variable and the load is fixed, it will be seen by inspection that (33) or (37) will become larger as Z_1 in the denominator is diminished in magnitude. Consequently, if we have a given load, such as a loud-speaker, and have the choice of a number of amplifying tubes to supply energy to the load, all of which have the same value of μ (*i.e.*, the same e.m.f.), the tube having the lowest internal resistance will supply the greatest amount of power for a given value of grid voltage. This fact is frequently overlooked in the desire to "match impedances" on any and all occasions. The impedances of the load and the source should be made equal only when the load is the variable.

The above expressions are applicable to a vacuum tube as a source of e.m.f., provided the internal resistance and amplification factor of the tube remain constant. This is approximately true if the signal voltage applied to the grid is small. For signal voltages of large amplitude, r_p will change throughout the cycle and therefore is some function of e_g. Consequently, the above derivations are no longer valid, as the internal resistance was treated as a constant when the expression for the power was differentiated. The conditions for maximum power output for large signal voltages will be considered presently.

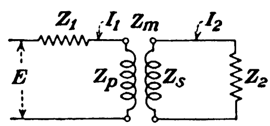


FIG. 117.—Load Z_2 coupled to a source of e.m.f. by means of a transformer.

When the impedance of the source and the load are both fixed and are unequal in magnitude, a transformer may be inserted between them to step up or step down the impedance of the load for maximum transfer of power. The secondary current in Fig. 117 is

$$I_2 = \frac{-EZ_m}{(Z_1 + Z_p)(Z_2 + Z_s) - Z_m^2} \tag{39}$$

Assuming an ideal transformer whose primary and secondary impedances are infinite and letting $a^2 = Z_s/Z_p$ $(a = N_s/N_p)$, then $Z_s = a^2Z_p$ and $Z_m = \sqrt{Z_pZ_s} = aZ_p$. Inserting these relations in (39) and neglecting finite quantities in comparison with infinite quantities, we get

$$I_2 = \frac{-EaZ_p}{Z_p(a^2Z_1 + Z_2)} = \frac{-Ea}{a^2(R_1 + jX_1) + (R_2 + jX_2)} \tag{40}$$

The absolute magnitude of (40) is

$$|I|_2 = \frac{Ea}{\sqrt{(a^2R_1 + R_2)^2 + (a^2X_1 + X_2)^2}} \tag{41}$$

If a is the variable, $|I_2|$ is a maximum when

$$a^2 = \sqrt{\frac{R_2^2 + X_2^2}{R_1^2 + X_1^2}} = \frac{|Z_2|}{|Z_1|} \quad \text{or} \quad a = \sqrt{\frac{|Z_2|}{|Z_1|}} \tag{42}$$

Therefore the power in Z_2 of Fig. 117 will be a maximum when the ratio of transformation of the transformer is made equal to the square root of the impedance ratio of the load and the source. The transformer winding having the greater number of turns is connected to the larger of the two impedances. While this relation assumes the use of an ideal transformer, it can be applied to an actual transformer with sufficient accuracy for most purposes.

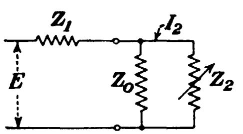


FIG. 118.—Source of e.m.f. with a fixed and a variable load shunted across it.

The source of e.m.f. may occasionally have a fixed value of impedance Z_0 permanently shunted across it, as in Fig. 118, and also shunted across the source is an adjustable load Z_2. What must be the value of this load impedance so that the power absorbed by Z_2 will be a maximum? For simplicity we will assume that the various impedances are all pure resistances. The current in Z_2 is given by

$$I_2 = \frac{ER_0}{R_1R_0 + R_2R_0 + R_1R_2} \tag{43}$$

The power in R_2 is

$$P_2 = \frac{E^2R_0^2R_2}{(R_1R_0 + R_2R_0 + R_1R_2)^2} \tag{44}$$

Differentiating (44), we find that the power in R_2 will be a maximum when

$$R_2 = \frac{R_1R_0}{R_1 + R_0} \tag{45}$$

The right-hand member of (45) will be recognized as the resultant resistance of R_1 and R_0 when connected in parallel. Therefore the circuit of Fig. 118 is equivalent to a source whose

e.m.f. is E and whose internal resistance is R_1 and R_0 in parallel. It can be shown in a similar manner for impedances that

$$|Z_2| = \left|\frac{Z_1 Z_0}{Z_1 + Z_0}\right| \tag{46}$$

for the case when $Z_2 = R_2 + jQR_2$.

71. Power Amplifiers.—The various circuits used with voltage amplifiers are also used as power amplifiers. Triodes for this purpose when used as Class A amplifiers usually have a relatively low value of r_p in order to obtain a reasonable amount of power output with a nominal value of plate-supply voltage. This results in a comparatively low value of μ so that a large grid-excitation voltage must be used. In recent years pentodes have been used to a considerable extent, as these tubes require a much

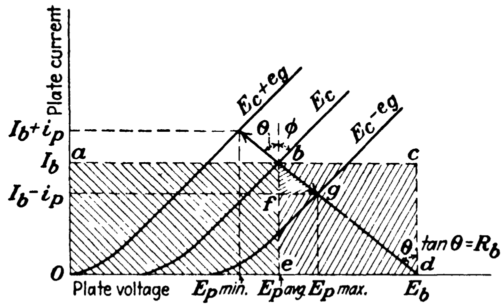


Fig. 119.—Areas representing power in the plate circuit.

smaller value of signal voltage for a given output. These will be discussed later.

The amount of power amplification under Class A operation is theoretically infinite as the grid is at all times negative with respect to the filament so as to prevent the flow of grid current. Actually, the input impedance of the tube is not infinite, owing to the input capacitance and conductance, the latter being caused by dielectric losses and reflected plate load. This input impedance is usually high enough to be disregarded from a power viewpoint at audio frequencies, except as it affects the performance of the preceding voltage amplifier. The energy represented by the useful output of the power tube is furnished by the plate-supply voltage E_b, as the term μe_g is a fictitious e.m.f.

The power relations in a triode for a pure resistance load R_b in the plate circuit may be shown by means of the I_p-E_p charac-

teristics of the tube as shown in Fig. 119. The power furnished by the B-supply will be E_bI_b, which is equal to the area $Oacd$. The average power supplied to the tube is $E_{p_{avg}}I_b$ and is equal to the area $Oabe$. The average power lost in the resistance R_b is $(E_b - E_{p_{avg}})I_b$ and is equal to the area $ebcd$. The average value of the useful or alternating-current power will be

$$P_{ac} = \frac{E_{p_{max}} - E_{p_{avg}}}{\sqrt{2}} \times \frac{i_p}{\sqrt{2}} = \frac{e_p i_p}{2}$$

since e_p and i_p are the maximum values of the alternating quantities. This is evidently equal to the area of the triangle fbg. Since $\tan\theta = R_b$ and $\tan\phi = r_p$, the useful power will be a maximum when $\theta = \phi$, assuming e_g to be small.

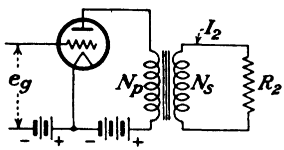


FIG. 120.—Load coupled to a triode by means of a transformer.

The usual method of coupling the load to the tube is by means of an output transformer of suitable ratio, as shown in Fig. 120. The circuit of Fig. 107 is also used, but it requires a load impedance which is comparable to r_p of the tube. The use of a transformer removes this restriction on the value of the load impedance, as the ratio of transformation can always be chosen so that impedance "looking into" the primary of the transformer will be of the proper value relative to r_p.

The power in R_2 of Fig. 120 will be, assuming an ideal transformer,

$$P = I_2^2R_2 = \left(\frac{1}{a}i_p\right)^2 R_2 = \frac{\mu^2 e_g^2 R_2}{a^2\left(r_p + \frac{1}{a^2}R_2\right)^2} = \frac{\mu^2 e_g^2 a^2 R_2}{(a^2 r_p + R_2)^2} \tag{47}$$

where $a = N_s/N_p$. If the load is an impedance $Z_2 = R_2 + jX_2$, (47) becomes

$$P = \frac{\mu^2 e_g^2 a^2 R_2}{(a^2 r_p + R_2)^2 + X_2^2} \tag{48}$$

The power will be either the maximum value or the average value, depending on whether e_g is the maximum or effective value of the signal voltage.

These expressions for the power again assume that μ and r_p are constant which is true only for small values of e_g. Power

amplifiers will ordinarily have a comparatively large value of e_g impressed upon them, so that graphical methods of determining the power output are usually employed, especially since the construction enables the amount of amplitude distortion to be determined at the same time.

72. Graphical Determination of Power Output and Distortion. Since the I_p-E_p characteristics are not parallel straight lines, there will be a certain amount of amplitude distortion present. In triodes the distortion is chiefly a second harmonic of the

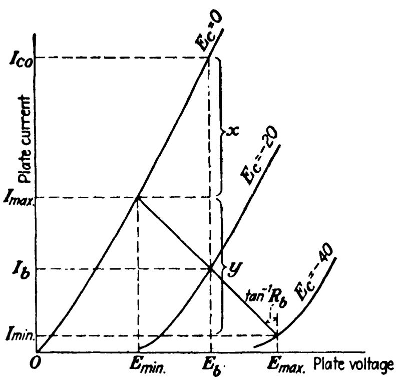


FIG. 121.—Construction for maximum undistorted power output.

signal voltage, assuming the latter to be composed of a single frequency so that the sum and difference frequencies discussed in Sec. 60 will be absent. This second harmonic causes the current and voltage in the plate circuit to be unsymmetrical about the time axes along I_b and E_p. The amplitude of the fundamental can be obtained by taking one-half of the difference of the maximum and minimum values, which eliminates the second harmonic.

The average power output at fundamental frequency, from Fig. 121 will be

$$P = \frac{E_{max} - E_{min}}{2\sqrt{2}} \times \frac{I_{max} - I_{min}}{2\sqrt{2}} = \frac{1}{8}(E_{max} - E_{min})(I_{max} - I_{min}) \tag{49}$$

This diagram is for the circuit of Fig. 120 and assumes an ideal transformer. The load line makes an angle of $\tan^{-1} R_b$ with the vertical, where $R_b = R_2/a^2$. The average plate voltage will

coincide with E_b since there is no resistance drop due to the flow of the direct current I_b in the transformer primary. The construction for the case of an actual transformer with the primary voltage drop taken into account is the same as Fig. 108, where $E_b - E_{pavg}$ is equal to this voltage drop.

The distortion is greatly increased if the minimum value of plate current under operating conditions is allowed to fall to too low a value, as the curvature in the characteristic is pronounced in this region. If the minimum current in Fig. 121 is fixed by distortion considerations and the path of operation is bounded on the other end by the curve $E_c = 0$, the value of R_b for maximum power output may be found as follows:

$$R_b = \frac{E_{\max} - E_{\min}}{I_{\max} - I_{\min}} \tag{50}$$

$$r_p = \frac{\frac{1}{2}(E_{\max} - E_{\min})}{I_{co} - I_{\max}} \tag{51}$$

Substituting the value of $(E_{\max} - E_{\min})$ from (51) in (49)

$$P = \frac{r_p}{4}(I_{co} - I_{\max})(I_{\max} - I_{\min}) = \frac{r_p}{4}xy \tag{52}$$

The distance $x + y$ is constant, as $I_{\min}$ and I_{co} are fixed, so that

$$x + y = c \qquad \text{and} \qquad P = \frac{r_p}{4}xy$$

or

$$P = \frac{r_p}{4}x(c - x) = \frac{r_p}{4}(cx - x^2)$$

If P is to be made a maximum, $dP/dx = 0$, which gives us $c = 2x$, so that $x = y$. Therefore

$$(I_{co} - I_{\max}) = (I_{\max} - I_{\min})$$

From (50)

$$I_{\max} - I_{\min} = \frac{E_{\max} - E_{\min}}{R_b}$$

and from (51)

$$I_{co} - I_{\max} = \frac{E_{\max} - E_{\min}}{2r_p}$$

so that

$$R_b = 2r_p \tag{53}$$

This has been termed the condition for maximum undistorted output, so designated because it results in the maximum power that can be obtained under the conditions that were imposed on the problem.[1] Values of load resistance higher than this will cause some reduction in output, but the amount of distortion will be reduced by a greater percentage, as will be seen in Fig. 123. It may happen that when $R_b = 2r_p$, the distortion may be more

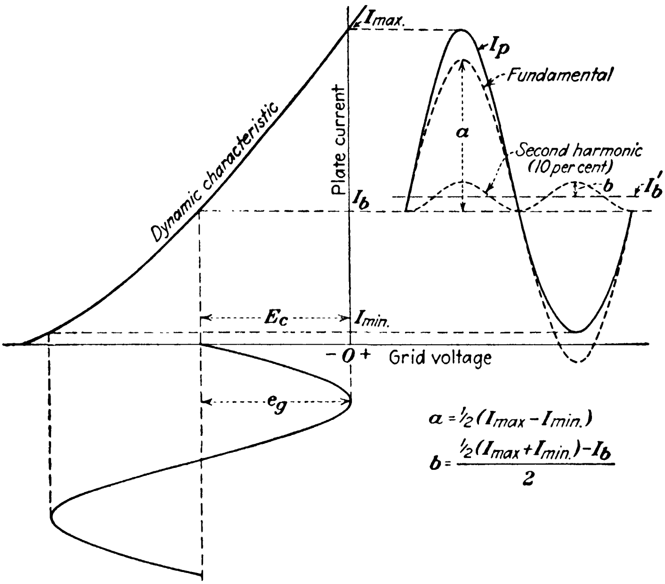


FIG. 122.—Production of second harmonic in a triode.

than is desirable, so that at the present time maximum undistorted output is taken to mean the maximum power output into a resistance load under the conditions that the total generated harmonics shall not exceed 5 per cent.

The amount of distortion can be determined from the maximum and minimum plate currents in Fig. 121, assuming that harmonics higher than the second are negligible. Figure 122 shows the dynamic I_p-E_g characteristic of a triode with a sine wave of voltage impressed on the grid. The resultant plate-current wave is I_p and the respective amplitudes of the funda-

[1] For examples of calculations under various imposed conditions see J. C. Warner and A. V. Loughren, The Output Characteristics of Amplifier Tubes, *Proc. I.R.E.*, vol. 14, p. 735, December, 1926.

mental and second harmonic are a and b. From the figure

$$I_{\max} = I_b + a + 2b \tag{54}$$
$$I_{\min} = I_b - a + 2b \tag{55}$$

Subtracting,

$$a = \tfrac{1}{2}(I_{\max} - I_{\min}) \tag{56}$$

Adding (54) and (55) and solving for b

$$b = \frac{\tfrac{1}{2}(I_{\max} + I_{\min}) - I_b}{2} \tag{57}$$

The percentage of second harmonic in terms of the fundamental is

$$\%\ \text{2nd} = 100\frac{b}{a} = 100\frac{\tfrac{1}{2}(I_{\max} + I_{\min}) - I_b}{I_{\max} - I_{\min}} \tag{58}$$

The curvature of the characteristic causes partial rectification to take place and the average value of plate current rises from I_b to I_b' when the signal voltage e_g is impressed on the grid. It will be observed that this increase is equal to the amplitude b of the second harmonic. This rise in plate current when distortion occurs can be readily observed by means of a suitable direct-current milliammeter in the plate circuit. With no amplitude distortion present, the needle of the meter should remain stationary while the signal voltage is being impressed. The amount of distortion present can be gaged by the increases in plate current as the signal fluctuates in intensity.

The increase in plate current due to rectification causes the voltage drop in the primary winding of the output transformer to increase from I_bR to $I_b'R$ when the signal voltage is applied. This must be corrected for, if appreciable, by shifting the load line upward so that the operating point O' in Fig. 108 coincides with I_b' instead[2] of I_b.

The variation of power output and second-harmonic distortion is shown in Fig. 123 for a type 45 triode. The values were determined from the I_p-E_p characteristics of the tube by means of (49) and (58). As will be noted the power output is a maximum for a load of 3400 ohms, which checks fairly well with the relation given by (53). From the standpoint of distortion it would be advisable to have the load resistance about 4000 or

[2] C. E. KILGOUR, *Proc. I.R.E.*, vol. 19, p. 42, January, 1931.

5000 ohms with this particular tube. This would reduce the power output only about 5 per cent for a given signal voltage, but the distortion would be reduced about 50 per cent.

73. Approximate Determination of the Output of a Triode.—The power output of a triode with a resistance load in the plate circuit will be approximately

$$P = \frac{\mu^2 e_g^2 R_b}{2(r_p + R_b)^2} \tag{59}$$

This expression assumes the I_p-E_p characteristics to be parallel straight lines and will be the average power if e_g is the

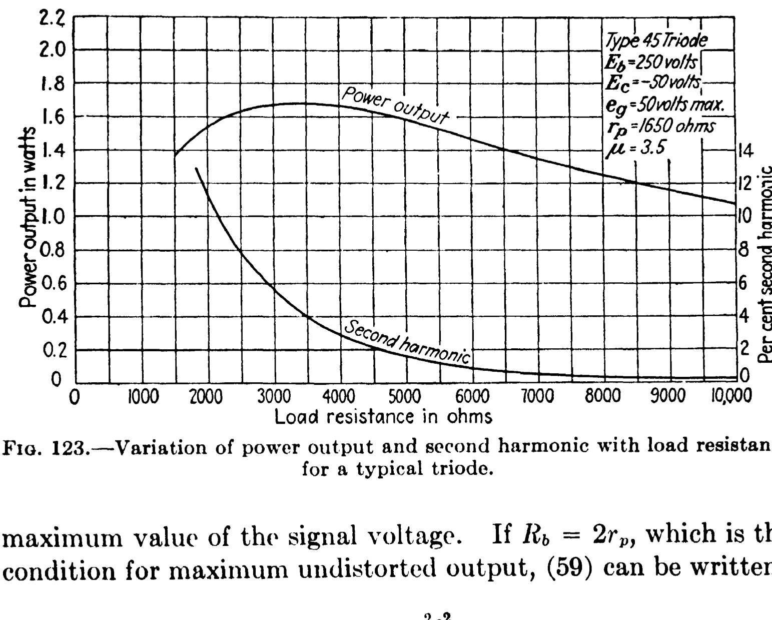


FIG. 123.—Variation of power output and second harmonic with load resistance for a typical triode.

maximum value of the signal voltage. If $R_b = 2r_p$, which is the condition for maximum undistorted output, (59) can be written

$$P = \frac{\mu^2 e_g^2}{9r_p} \tag{60}$$

This expression is useful in estimating the power output of a tube for a given signal voltage. For Class A operation the grid is negatively biased so that the maximum signal voltage causes the potential of the grid with respect to the filament to vary from zero to an amount somewhat less than cut-off. Assuming the cut-off voltage of the grid to be given by E_b/μ, the maximum permissible value of e_g will be $E_b/2\mu$. Substituting this value of

e_g in (60), the maximum value of the average power output will be approximately

$$P = \frac{E_b^2}{36r_p} \tag{61}$$

The peak value of power output will be twice this value. This expression is useful in roughly estimating the maximum power output of a triode under Class A operation in terms of its plate-supply voltage and plate resistance. Where a relatively large amount of power is required with a moderate plate voltage it is evident from (61) that the internal resistance of the triode must be small.

74. Push-pull Amplifiers.—There will always be more or less amplitude distortion in the preceding types of power amplifiers, owing to the curvature of the characteristic. A large value of load resistance will make the dynamic characteristic more nearly linear, but this entails some sacrifice in power output. The arrangement of Fig. 124, known as a push-pull circuit, can be made to reduce amplitude distortion to negligible proportions. The input transformer has a center tap brought out from the secondary winding which is connected through the C battery to the common filament connection of the two tubes. Assuming an alternating signal voltage to be induced in the transformer secondary, the potential of one grid will become more positive with respect to its filament while the grid of the other tube will become increasingly negative. This will cause the plate current of the first tube to increase while the plate current of the other tube diminishes. The alternating voltages across the two halves of the primary of the output transformer are in phase, although the direct-current voltage drops are in opposition. The net magnetomotive force exerted by the two halves of the primary is zero when no signal is impressed, as I_{b1} and I_{b2} are flowing in opposite directions around the primary. With a signal voltage impressed, the magnetomotive force produced by the plate current of the two tubes is equal to that which would be produced by a current equal to the difference between

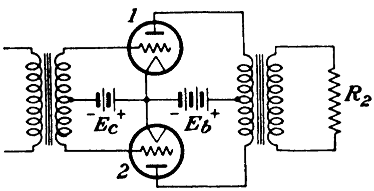


FIG. 124.—Push-pull amplifier circuit.

the current flowing in tube 1 and that flowing in tube 2, flowing in one-half of the primary winding.

This is shown in the diagram of Fig. 125. The dynamic I_p-E_g characteristics of the two tubes are drawn with those of

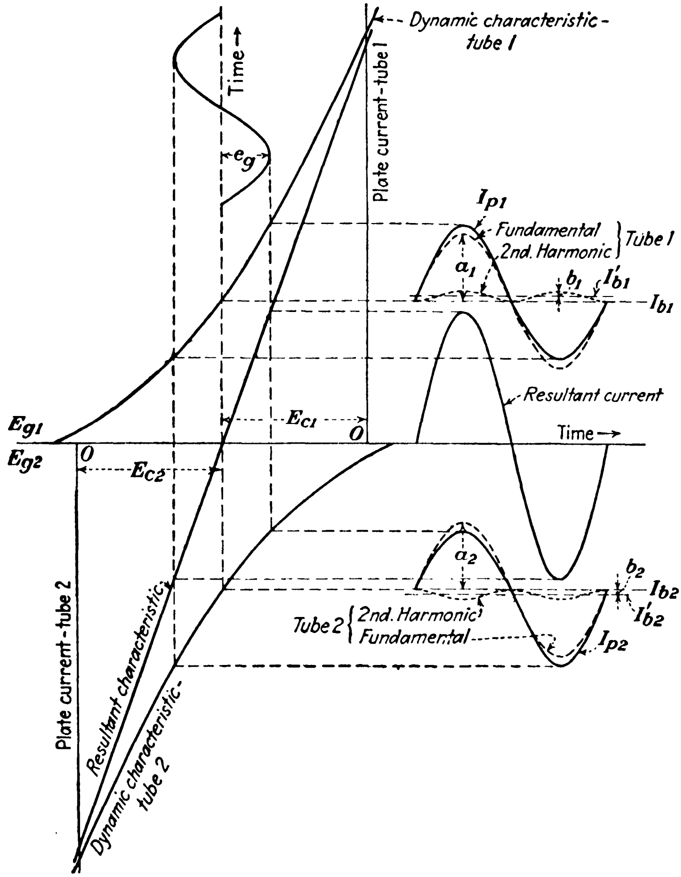


FIG. 125.—Graphical determination of the resultant characteristic of a push-pull amplifier showing the cancellation of even harmonics produced by the curvature of the individual tube characteristics.

tube 2 inverted, so that zero grid voltage is at the left-hand end of the diagram and the ordinates of plate current are drawn in the negative direction. The resultant characteristic is obtained by taking the difference between the two plate currents for corresponding grid voltages. It is interesting to observe that this resultant is a straight line in spite of the considerable curva-

ture present in the two individual dynamic characteristics. This is not always the case, although the resultant will always be straighter than the two individual characteristics. The individual plate currents and their constituent harmonics, obtained in the manner of Fig. 122, will be as shown. As will be seen, the second harmonics are in opposition and cancel. All other even harmonics will also be in phase opposition, so that the resultant current will contain no even harmonics, assuming

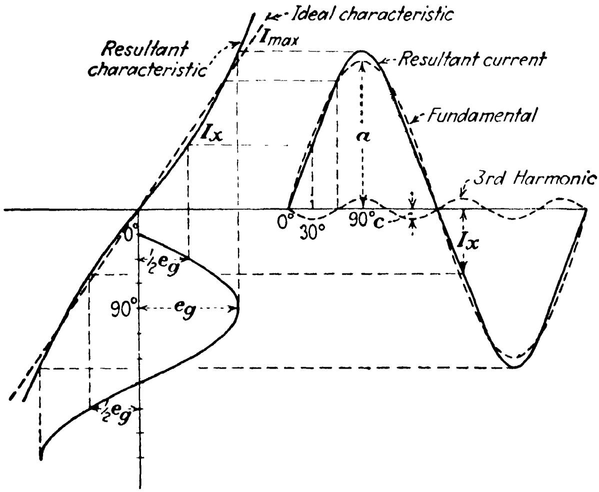


FIG. 126.—Production of a third harmonic due to curvature in the resultant characteristic of a push-pull amplifier.

identical tubes. Odd harmonics will be in phase and hence will not cancel.

The resultant current shown is that value of current which would have to flow through one-half of the primary winding in order to produce the same effect as the difference between the two plate currents flowing through the whole primary. This current is equal to $a_1 + a_2$ and can be obtained by projecting e_g on the resultant characteristic. The voltage e_g is the signal voltage applied to the grid of either tube and is equal to one-half of the secondary voltage of the input transformer.

When the dynamic characteristics of the individual tubes are such as to produce odd harmonics in their plate-current wave shapes, the resultant characteristic will deviate from a straight line, as shown in Fig. 126. The resultant current will then

contain odd harmonics, the third being the most important. At overloads an appreciable amount of fifth harmonic may be produced. The magnitude of the third-harmonic distortion can be determined from the maximum value of the resultant current and an ordinate I_x corresponding to $\frac{1}{2}e_g$, as follows: From Fig. 126

$$I_{max} = a + c \tag{62}$$

$$I_x = \frac{a}{2} - c \tag{63}$$

Adding (62) and (63), we find the amplitude of the fundamental to be

$$a = \frac{I_{max} + I_x}{1.5} \tag{64}$$

Subtracting (63) from (62) and solving for c we get

$$c = \frac{I_{max} - 2I_x}{3} \tag{65}$$

The percentage of third harmonic present in terms of the fundamental is

$$\% \text{ 3rd} = 100\frac{c}{a} = 100\,\frac{I_{max} - 2I_x}{2(I_{max} + I_x)} \tag{66}$$

This expression assumes that the fifth harmonic is negligible.

The dynamic I_p-E_g characteristics have to be constructed for each particular value of load resistance used, so that a construction due to B. J. Thompson[3] which obtains a family of composite characteristics for a push-pull amplifier directly from the I_p-E_p characteristics of the tube will reduce the amount of labor involved. This construction is similar to that of Fig. 125 and is shown in Fig. 127 for a typical triode. The curves for tube 2 are inverted as before, with zero plate voltage at the right-hand end of the diagram and the ordinates of plate current drawn in the negative direction. The two plate-voltage scales will coincide at the point E_b. The resultant or composite characteristics are shown by the heavy lines and are obtained by taking the difference between the two plate currents for corre-

[3] *Proc. I.R.E.*, vol. 21, p. 591, April, 1933.

sponding grid voltages. It will be noted that these composite characteristics are very nearly parallel straight lines in this particular case. The load line is drawn in the ordinary manner, making an angle with the vertical whose tangent is R_b, and passing through the point E_b at zero-resultant plate current. In the case illustrated, $E_b = 250$ volts and $R_b = 1000$ ohms. This

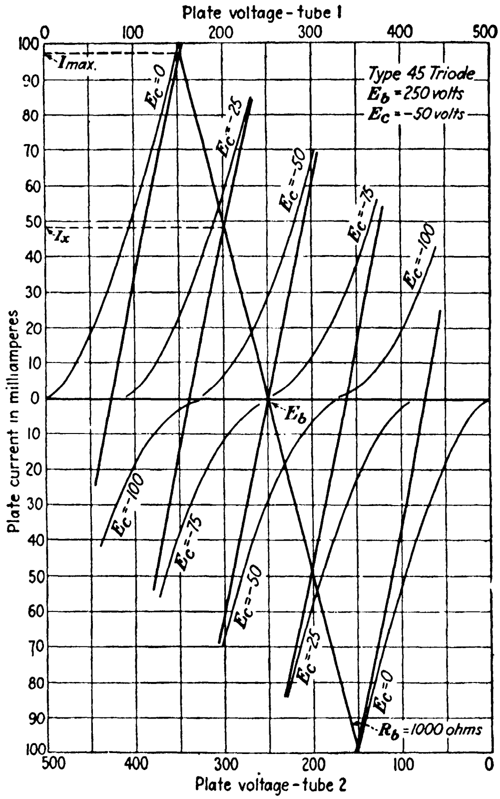


FIG. 127.—Composite characteristics of a push-pull amplifier.

value of R_b is the impedance looking into one-half of the primary winding of the output transformer. The plate-to-plate load resistance is four times this value, or 4000 ohms, which would be the value of R_2 in Fig. 124, assuming the total number of turns on the primary and secondary to be equal.

The total power output of the amplifier is readily obtained from

$$P = \left(\frac{I_{\max}}{\sqrt{2}}\right)^2 R_b = \frac{I^2_{\max} R_b}{2} \tag{67}$$

assuming the odd harmonics to be negligible. If they are sufficient to produce appreciable error, the output at fundamental frequency may be obtained from

$$P = \frac{a^2 R_b}{2} = \frac{(I_{\max} + I_x)^2 R_b}{4.5} \tag{68}$$

where a is the amplitude of the fundamental and is given by (64).

The slopes of the composite characteristics are seen to be practically constant, so that the apparent internal resistance of the source is also constant. This is due to the coupling between the two tubes through the two halves of the output-transformer primary. Hence, maximum power output will be obtained in this case when the resistances of the load and the source are made equal. This optimum value of load resistance sometimes results in values of $I_{\max}$ which are excessive and the resultant plate loss exceeds the allowable rating for the tube. In such cases a load resistance somewhat greater than the optimum value must be used. The power dissipated at the plate in the form of heat will be the power supplied by E_b minus the output to the load, including the losses in the output transformer.

The value of the apparent internal resistance is obtained by taking the cotangent of the slope of the composite characteristics. The value of this resistance for the type 45 tubes operating under the conditions of Fig. 127 is 800 ohms, which is about one-half r_p for this tube at $E_b = 250$ volts and $E_c = -50$ volts. The power output and percentage distortion in the form of the third harmonic are shown in Fig. 128, for various values of plate-to-plate load resistance. The values of the latter are four times R_b, since the total output of the amplifier is thought of as being impressed across only one-half of the primary winding.

Comparing Fig. 128 with Fig. 123 for a single tube of the same type, with the same signal, bias, and plate-supply voltages, it is seen that the push-pull amplifier is capable of delivering almost three times as much power as the single tube and with considerably less amplitude distortion. The erroneous statement is frequently made that the output of two tubes in push-pull is theoretically only twice that of a single tube, and that slightly

greater output than this can be obtained by impressing somewhat larger signal voltages than could be tolerated with a single tube without excessive distortion. The greater output of the push-pull circuit is due to the lower apparent internal resistance of the combination, which may be viewed as though it were a single

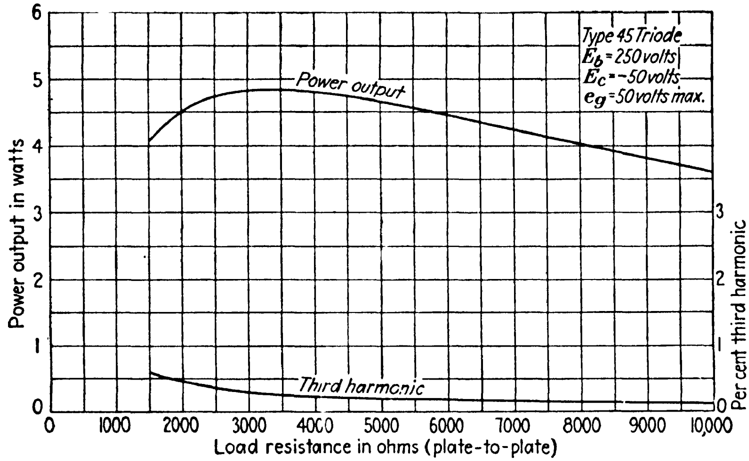


FIG. 128.—Variation of power output and third harmonic with load resistance for two typical triodes in push-pull.

tube of internal resistance R_o and load resistance R_b. The average power output can be computed from

$$P = \frac{\mu^2 e_g^2 R_b}{2(R_o + R_b)^2} \tag{69}$$

where e_g is maximum value of the signal voltage applied to one grid, and R_o is given by $\Delta E_p/\Delta I_p$ of the composite characteristics. The value of R_o is approached by r_p of the tube for high values of plate current. For approximate estimates of output using (69), R_o may be taken as $r_p/2$.

The maximum power output will occur in push-pull amplifiers operating Class A when $R_b = R_o$ and (69) becomes

$$P_{\max} = \frac{\mu^2 e_g^2}{8R_o} \tag{70}$$

75. Advantages of Push-pull Amplifiers.—In addition to the considerable reduction in amplitude distortion and the increased power output per tube, the push-pull circuit has other advan-

tages. The plate supply for most power amplifiers is obtained from rectified alternating current which may contain more or less ripple, depending upon how effectively the rectifier output is filtered. In push-pull circuits the small alternating current in the plate circuit produced by this ripple voltage flows in opposite directions in the two halves of the output-transformer primary, so that no voltage is induced in the secondary. This assumes that the two tubes are identical and that the impedances of the two halves of the primary winding are accurately balanced.

The biasing voltage E_c is usually obtained by using the IR drop across a resistance inserted in the negative lead of the plate supply. In the ordinary type of amplifier any unfiltered ripple flowing through this resistance is introduced into the grid circuit and is amplified by the tube. The amplified signal current will also flow through this resistance and will be fed back into the input circuit in opposition to the signal voltage unless suitable filtering circuits are provided. In push-pull amplifiers the ripple voltage superimposed on E_c affects both grids alike, so that the plate currents resulting from this source annul each other in the output transformer. No signal currents other than even harmonics of the signal frequency will flow in the B supply circuit, so that filtering circuits across the biasing resistance to avoid feed-back difficulties are not required in push-pull circuits operating as Class A amplifiers.

When alternating current is used for heating the filament, the electron stream may be modulated by both the alternating electrostatic and electromagnetic fields surrounding the filament, which is equivalent to a small alternating potential applied to the grid. The external connections to the filament are made to a center tap in the secondary winding of the filament heating transformer, or else to the mid-point of a low resistance shunted across the filament. If this connection is displaced from the electrical center, an alternating potential will again be applied to the grid. Here again, all these effects tend to cancel in the push-pull circuit. Tubes using indirectly heated cathodes are relatively free from hum caused by the use of alternating current for heating purposes unless followed by an amplifier having a high gain. Here the ordinary construction used in heater-type tubes is not sufficiently free from hum to be used in the early stages of high-gain amplifiers and specially designed tubes have

to be employed.[4] The majority of power amplifiers work directly into the useful load without subsequent amplification and filament-type tubes can be used without objectionable hum.

It has already been mentioned that in push-pull circuits the direct-current components of plate current of the two tubes flow in opposite directions in the two halves of the output-transformer primary, so that the core is not subjected to a continuous magnetizing force. This greatly increases the incremental permeability of the core and enables a higher primary inductance to be secured with the same number of turns. This feature enables core materials of various magnetic alloys to be used, for the high incremental permeabilities of these materials fall

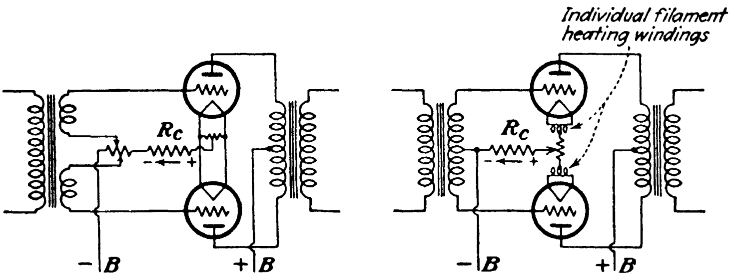


FIG. 129.—Methods of adjusting individual biasing voltages in a push-pull amplifier for equal plate currents.

very rapidly as the continuous magnetizing force is increased, as discussed in Chap. III. When alloy cores are used in the output transformers of push-pull amplifiers, it is important that the two tubes have the same values of I_b. As it is inconvenient to sort through a group of tubes to find a pair that are alike in this respect, provision is sometimes made for the bias voltage of one or both of the tubes to be adjusted individually for equal plate currents, as shown in Fig. 129. This method of securing the required C bias is discussed in the following section.

76. Methods of Obtaining C Bias.—With battery operation of receiving sets the negative bias for the grids of the various tubes is readily obtained by the use of a C battery for this purpose. In sets operated entirely from an alternating-current lighting circuit, the biasing voltage is obtained from the voltage drop across a suitable resistance in series with or shunted across the

[4] J. O. McNALLY, A "Low-hum" Vacuum Tube, *Bell Lab. Rec.*, vol. 11, p. 158, February, 1933.

plate-voltage supply. In the radio-frequency amplifying stages the bias for all tubes can be obtained from a resistance common to the plate circuits of all tubes, as feed-back difficulties due to this common coupling between stages are easily avoided by the use of by-pass condensers. Audio-frequency amplifiers are usually biased by individual resistances, as the much lower values of frequency dealt with makes it difficult to properly by-pass these currents when a common resistor is employed, unless rather large condensers are used.

When a resistance R_c is inserted in the negative lead of the supply voltage, as shown in Fig. 130, the average voltage drop across this resistance will be I_bR_c, and if the grid return lead is connected to the lower end of R_c, the grid will be held negative with respect to the filament by the amount of this drop. The alternating component of plate current i_p will also flow through R_c and will produce a voltage drop i_pR_c which will be 180 degrees out of phase with the signal voltage e_s, if Z_b is a pure resistance R_b. The alternating voltage applied to the grid will be

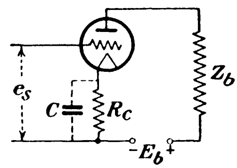


FIG. 130.—Self-biased amplifier.

$$e_g = e_s - i_pR_c \tag{71}$$

and the plate current will be

$$\begin{aligned} i_p &= \frac{\mu(e_s - i_pR_c)}{r_p + R_b + R_c} \\ &= \frac{\mu e_s}{r_p + R_c(\mu + 1) + R_b} \end{aligned} \tag{72}$$

The apparent effect of this biasing resistance, in the absence of the by-pass condenser C, is to augment the internal resistance r_p of the tube by a factor $R_c(\mu + 1)$, which seriously reduces the power output. In the case of the type 45 tube previously considered, R_c would have to be about 1600 ohms to secure a negative bias of 50 volts, and with an amplification factor of 3.5, the term $R_c(\mu + 1)$ would be 7200, or about 4.5 times the value of r_p for this tube. In addition to this reduction in power output, there would also be considerable frequency distortion if the load were not a pure resistance. Assume that Z_b contains an appreciable amount of inductive reactance. Then at the higher frequencies i_p will be smaller than at lower frequencies for the same

value of e_s. Consequently, from (71), the net voltage impressed on the grid will diminish as the frequency of the impressed signal is reduced. Vector expressions must be substituted in (71) and (72) if reactance is present in the circuit.

If the biasing resistance is shunted by a condenser whose reactance is small compared with R_c at the lowest frequency with which we are concerned, the impedance offered by the combination to alternating currents of higher frequencies becomes progressively smaller, and the troubles caused by the reversed feed-back or "degeneration" will be eliminated. This requires a rather sizable condenser. For example, the reactance of a 2-μf condenser at 60 cycles is 1326 ohms, which would still be

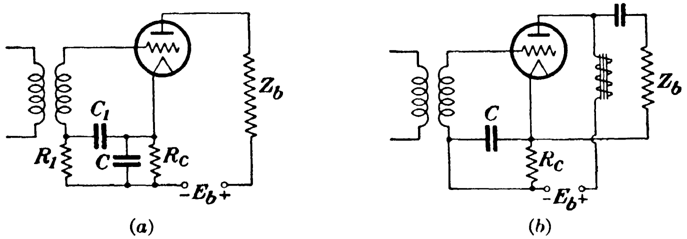


FIG. 131.—Methods of avoiding reversed feed-back in self-biased amplifiers.

comparable to the value of R_c, so that a much larger value of capacitance must be employed if undue discrimination against the low frequencies is to be avoided. The value of the plate current when a by-pass condenser is placed across R_c can be determined from (72) if the term R_c in the denominator is replaced by the vector expression for R_c and C in parallel.

One method which does away with reversed feed-back without the need of a very large by-pass condenser is the circuit illustrated in Fig. 131*a*. The bias is applied to the grid through a high resistance R_1, while the signal voltage is applied to the tube through the condenser C_1, the reactance of which is small compared to R_1. The high value of R_1 tends to prevent the feed-back voltage from entering the input circuit. The method of parallel feed of Fig. 131*b* attacks the problem in another way in that the alternating component of plate current does not flow through the biasing impedance. The push-pull circuit when operated Class A is free from feed-back troubles due to the biasing resistance. Furthermore, since two tubes are employed the value of R_c

required will be only one-half that needed by a single tube, as in the push-pull case the current flowing through R_c will be $2I_b$.

In all these self-biasing circuits the biasing voltage rises somewhat as the value of e_s increases. This is due to partial rectification in the tube which causes the average value of the plate current to rise from I_b to I'_b, as shown in Figs. 122 and 125.

77. Push-pull Amplifiers, Class B.—In order to secure greater outputs with a given tube and plate-supply voltage than can be

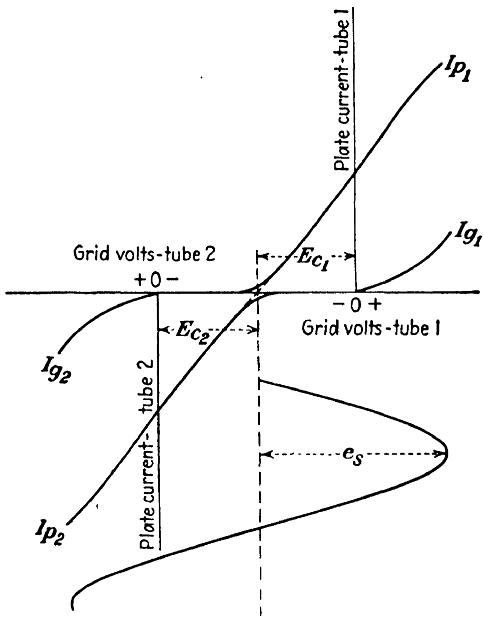


FIG. 132.—Plate- and grid-current characteristics of a push-pull Class B amplifier.

obtained under Class A operation, it is possible to operate a push-pull circuit as a Class B amplifier.[5] The circuit is the same as for Class A operation, except that the tubes are biased almost to cut-off, as shown in Fig. 132. By properly choosing the value of load impedance and bias voltage the resultant characteristic can be made to approximate a straight line. The dynamic characteristic must be essentially linear if distortion is to be avoided, as the curvature in the one characteristic is not offset by that of the other except in the vicinity of cut-off. In order

[5] L. E. BARTON, High Audio Power from Relatively Small Tubes, *Proc. I.R.E.*, vol. 19, p. 1131, July, 1931.

to obtain a large power output the grids are driven positive during a portion of the cycle, so that the preceding stage of amplification must be capable of furnishing the necessary power represented by the flow of grid current. In addition, the preceding amplifier—usually termed the "driver"—must also have good voltage regulation, or the tops of the signal voltage e_s will be flattened. Poor regulation is caused by high internal impedance in the source of e_s, so that a step-down ratio is usually used in the push-pull input transformer. The internal plate resistance of the driver tube will then be reduced by the square of the transformation ratio as viewed from the push-pull stage. The step-down ratio is limited by the maximum undistorted voltage that can be developed across the primary of the input transformer by the tube used in the preceding stage. The Class B stage is sometimes excited by a smaller push-pull Class A amplifier, which, in addition to its other advantages, will have an apparent value of plate resistance of approximately one-half that of a single tube, as was pointed out in connection with Fig. 127.

This mode of operation is a very satisfactory solution to the problem of obtaining reasonable power outputs for loud-speaker operation in battery-operated receiving sets. The cost of the plate and filament energy supplied to the last or power stage in the ordinary alternating-current receiving set is of little concern to the user. But in a set operated entirely from dry batteries the cost of the energy is about $10 per kilowatt-hour, so that the *B* battery requirements of the last audio stage, as to the current drain and the number of cells needed, are important considerations.

As an example of the relatively large amount of output obtainable under Class B operation, a pair of type 30 triodes requiring 2 volts at 0.06 amp. each for filament heating are capable of delivering from 1 to 2 watts output—depending on the distortion allowable—with only 157.5 volts on the plate and a grid bias of -15 volts. The total plate current varies from 1 ma when no signal is being received to a peak value of 50 ma. The fluctuations of plate current with the signal voltage are similar in appearance to the output of an unfiltered full-wave rectifier. The average plate current is very much less than with Class A operation because of the small value of I_p when no signal is being received. This results in much longer service life of *B* batteries

in battery-operated sets. This is particularly true in cases where the receiving set is turned on continuously but receives messages only at intervals, such as police radio cars. Likewise, when other *B*-supply devices are used which operate from a storage-battery source, there will be a corresponding reduction in the drain on the latter. With Class A operation the plate power supplied remains practically constant and does not fluctuate appreciably with the received signal.

The typical power tubes used in Class A applications are not well suited for Class B operation. Their amplification factors are low so that a large grid-excitation voltage is necessary, which requires a relatively large amount of power on the part of the driver to swing the grids positive. Tubes with a high amplification factor are therefore more satisfactory.

The fluctuations in the total plate-current input to the two tubes are at a frequency which is twice that of the impressed signal and the variation is from almost zero to a value limited by the allowable plate dissipation of the tube. This requires a source of plate voltage having good regulation, such as a battery, so as to prevent similar fluctuations in the terminal voltage of the supply, which would affect other tubes connected to the same source. It is rather difficult to secure the regulation needed from the ordinary diode rectifier, because of its large internal resistance. The hot-cathode mercury-vapor type with its constant internal-voltage drop is satisfactory, but in receiving sets it is apt to produce objectionable noise owing to radio-frequency disturbances within the tube itself unless the rectifier is shielded and provided with radio-frequency choke coils in series with the plate leads. This difficulty is not present in amplifiers for public address systems and similar applications, so that mercury-vapor rectifiers are satisfactory sources of Class B plate power in these fields.

Self-bias is not practical in Class B amplifiers owing to the extreme fluctuations in the plate current. Either a battery or a small auxiliary rectifier must be used to supply the biasing voltage. This drawback has led to the design of tubes having a very high amplification factor so that they can be operated at zero bias. Grid current will then flow during the entire cycle of the excitation voltage, but this disadvantage is offset to some extent by the fact that the input impedance remains more nearly

constant during the cycle. An interesting example of a tube for this purpose is the type 46 which employs a double grid. These are wound one within the other in the form of a flattened helix and each is brought out to its respective pin on the tube

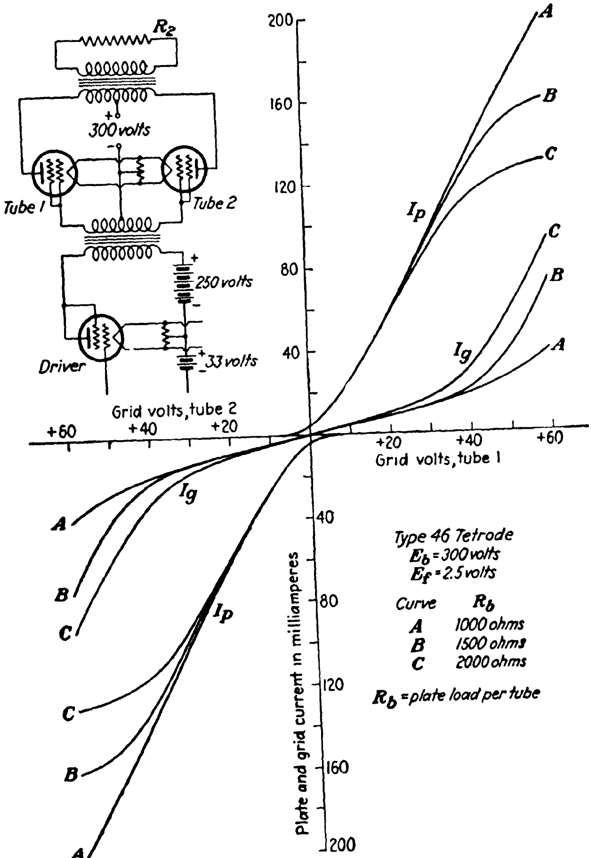


FIG. 133.—Dynamic characteristics for type 46 tetrodes used as push-pull Class B amplifiers.

base. For Class B operation both grids are connected together at the socket which results in a high amplification factor. The tube is also suitable for Class A applications, and when so used the outer grid is connected to the plate. This has the effect of moving the plate closer to the inner or control grid and reducing the distance b_2 in (33), Chap. VI, with the consequent reduction

in the amplification factor from about 65 with the first connection to 5.6 with the second. This possibility of dual application enables one tube to be used as a Class A driver of pair connected for Class B operation, as shown in the diagram of Fig. 133. The dynamic characteristics of the push-pull stage are shown in the same figure. This combination is capable of delivering an output of 16 watts into a load resistance R_b of 1300 ohms. The power required for grid excitation is 0.95 watt.

The load resistance R_b is the resistance of R_2 as viewed when looking into one-half of the primary of the output transformer. Thus, if the total number of turns on the primary and secondary of the output transformer in Fig. 133 are equal, the value of R_2 would have to be 5200 ohms in order for R_b to be 1300 ohms. The effective ratio of transformation of the output transformer is $N_2/\frac{1}{2}N_1$.

The input transformer alternately furnishes power to the grid of one output tube and then the other, so that its effective ratio of transformation would be $\frac{1}{2}N_2/N_1$ in computations involving the driver stage. This ratio of transformation should be chosen so that the load impedance R_b in the plate circuit of the driver tube is higher than the normal value for optimum output as a Class A power amplifier in order to insure low distortion in the driver stage. An allowance should be made for the transformer efficiency in computing the output of the driver stage. The usual transformer will have an efficiency approaching 80 per cent at output peaks.

Current flows in opposite halves of the primary of the output transformer during alternate cycles, so that mutual coupling effects between the two tubes are absent except at very low values of plate current. Consequently, the composite characteristics of Fig. 127 when applied to Class B operation will coincide with the static I_p-E_p characteristics of the tube, except for small values of plate current. The power output of a Class B stage can be determined graphically in the same manner as in Fig. 127 for Class A operation. The load line is drawn through E_b making an angle with the vertical of $\tan^{-1} R_b$. Since the power during one-half cycle is furnished by tube 1 and during the remaining half cycle by tube 2, the inverted characteristics for tube 2 in Fig. 127 do not need to be drawn. Equations (64) to (68) for Class A operation are all applicable to the Class B case. An

appreciable amount of fifth harmonic may be produced under certain conditions of operation and when this is the case, the values of third harmonic and power output given by (66) and (68) will become inaccurate.

Theoretically, no even harmonics are produced by push-pull Class B amplifiers, if the two tubes are identical and if the impedances of the two halves of the output transformer primary are balanced. Any dissimilarities in the characteristic curves of the two tubes will introduce even harmonics, and to a greater extent than in Class A operation. In the latter case the resultant characteristic is obtained as the difference between the two dynamic characteristics. This may be seen from Fig. 125. If the ordinates of the upper dynamic characteristic had all been increased by a fixed amount, the resultant characteristic would have been raised by this same amount without change of slope or curvature, and no even harmonics would have been introduced in the resultant current. But if these tubes had been operating Class B, even harmonics would have been produced by this change in the characteristic of tube 1. It is evident from this that the tubes used for Class B operation should be as closely matched as possible if the production of even harmonics is to be avoided. Even harmonics produced by the driver stage will, of course, be reproduced in the output of Class B stage.

The chief advantages of Class B operation are that large amounts of power can be obtained from moderate-sized tubes and that the plate power required is substantially less than with Class A for the same amount of useful output. The distortion is greater than with Class A operation.

Tubes combining two high-mu triodes contained in a single bulb and designed for Class B operation are on the market. The triode units have separate external terminals except for the cathode, which is common to both units. The two triodes may be connected in parallel to serve as a driver and in this way reduce the number of different types of tubes required.

A duplex tube of this type can also be used to produce a two-stage resistance-coupled amplifier as the high value of μ required for satisfactory Class B operation makes these tubes suitable for this purpose.

78. Push-pull Amplifiers, Class AB.—This mode of operation, sometimes called Class "A prime," is intermediate to Class A

and Class B. The grid bias employed is somewhat greater than for Class A but less than cut-off. The grids are not usually driven positive, although they may be if sufficient power is available in the preceding stage. Somewhat greater power output can be obtained than with Class A operation and with less amplitude distortion than with Class B. Self-bias can be used, but the average value of plate current fluctuates and the biasing resistance must therefore be shunted by a very large condenser or suitable filter network to minimize the resultant variations in the grid-bias voltage. Electrolytic condensers having a capacitance of 50 μf or larger are suitable for this purpose. An auxiliary rectifier is sometimes used for biasing, as in Class B operation. The use of a fixed bias of this sort enables somewhat larger outputs to be obtained than is possible with self-bias. The idle plate current and attendant dissipation of power may be made appreciably less than with Class A operation. This item is of importance in high-power modulators used in radio telephony. These modulators are merely large audio-frequency power amplifiers whose alternating-current output is made to vary the plate voltage supplied to a radio-frequency power amplifier, and in this way causes the radio-frequency output to vary in amplitude at an audio-frequency rate. In large transmitters the output of the modulator is rated in kilowatts rather than in watts, so that the attendant saving in the cost of plate power over what would be required with Class A operation is a factor worthy of consideration. Modulators can be also operated Class B.

Calculations of distortion and power output can be made by means of (66), (67), and (68). The composite characteristics are constructed in the same manner as in Fig. 127. For example, if E_b were made 275 volts and E_c was increased to -75 volts, the lower family of curves would all have to be displaced to the right by 50 volts and the point $E_b = 275$ will be the same for both groups. The composite characteristics can then be constructed from the following pairs of curves: $E_{c1} = -75$, $E_{c2} = -75$; $E_{c1} = -100$, $E_{c2} = -50$; etc.

Class AB operation produces somewhat less distortion than Class B and can be operated self-biased. The plate power consumed is somewhat greater, although less than with Class A.

79. Amplifiers Using Tetrodes.—The most common form of tetrode is the screen-grid type. This tube was developed pri-

marily for radio-frequency amplifiers, although it is used occasionally as a resistance-coupled voltage amplifier. The construction of this tube is shown diagrammatically in Fig. 134. The second or screen grid is introduced between the control grid and the plate, completely surrounding the latter so as to reduce, so far as possible, the capacitance between the control grid and plate. The magnitude of this capacitance is reduced in this manner to value less than 0.01 $\mu\mu$f. It will be shown later that the capacitance between control grid and plate serves electrostatically to couple the input and output circuits, enabling some of the amplified output energy to be fed back into the input circuit. A continuous oscillation may be sustained, depending upon the magnitude and phase of the energy thus fed back. The reactance of C_{gp} in the ordinary triode is usually too large to cause trouble from this source at audio frequencies.

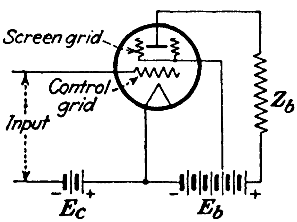


FIG. 134.—Connections of a screen-grid tube.

A positive potential of about one-half E_b, or less, is impressed on the screen grid. The I_p-E_p characteristics of a typical tube of this type are shown in Fig. 135 for a screen-grid potential of 90 volts. The values to the left of the line A are rather unstable, as the plate voltage is lower than that of the screen-grid in this region and the effects of secondary emission from the plate are pronounced. As the plate voltage increases from zero the plate current rises to a maximum and then begins to fall off. This reduction in the plate current is caused by the secondary electrons which are knocked out of the plate by the impact of the primary electrons, and which travel to the screen grid, since it is at a higher positive potential than the plate. The net plate current will be the algebraic sum of the electrons leaving, and arriving at, the plate. The number of secondary electrons which leave the plate may exceed the number which arrive, in which case the plate current will actually reverse and flow out of the plate. As the value of the plate potential approaches that of the screen grid, the number of secondary electrons which travel to the screen grid are reduced, as indicated by the reduction in the screen-grid current. The plate current abruptly rises and then

increases at a rather slow rate after its potential exceeds that of the screen grid. In this region the plate resistance is very high. The amplification factor is also large so that the mutual conductance is comparable to that of a triode of similar size. For the type 24-A tube shown in Fig. 135, $\mu = 400$ and $r_p = 400{,}000$ ohms at $E_p = 180$ volts and $E_c = -3$ volts.

The shape of the curve of grid current in the ordinary triode is often very similar to that of the plate current of the tetrode to

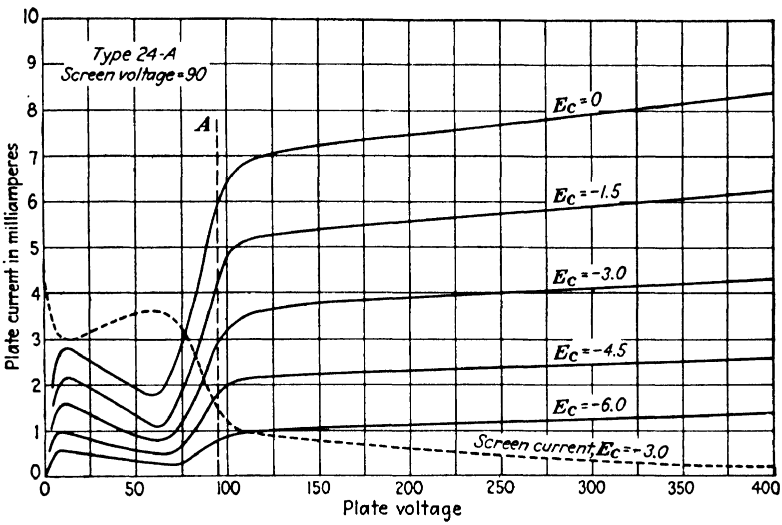


FIG. 135.—Characteristics of a typical screen-grid tetrode.

the left of the line A, and for the same reasons. When the grid is considerably positive, secondary emission occurs and these secondary electrons leave the grid and are attracted to the more positively charged plate. This causes the current in the grid to rise and then fall off as its positive potential is increased, in much the same manner as the plate current of a tetrode for low plate voltages. As the positive grid voltage approaches that of the plate the grid current begins abruptly to rise again as fewer secondary electrons are attracted away from the grid to the plate.

Tetrodes can be satisfactorily used as voltage amplifiers provided the amplitude of the impressed signal is not too large. They are not capable of delivering very much output when used

as Class A power amplifiers owing to the limitations imposed on the allowable plate-voltage swing. Serious distortion will result if the plate voltage falls below that of the screen grid. This limitation as to distortion is not a factor in radio-frequency power amplifiers, which usually operate Class C, so that screen-grid tubes are satisfactory in this field. They are designed to operate with a screen-grid voltage much lower in proportion to the plate voltage than is customary with receiving tubes.

80. Tetrodes with Space-charge Grid.—In this type of circuit the grid nearest the cathode is operated at a small positive potential, often by connecting it to the positive end of the filament, as shown in Fig. 136. This grid attracts electrons from the cathode and aids to some extent in reducing the space charge, which gives it its name. The second grid is used as the control grid.

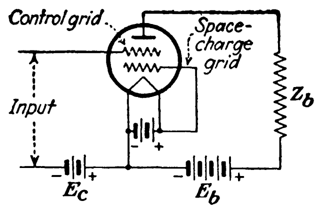


FIG. 136.—Connections of a tetrode with a space-charge grid.

The advantage of this circuit is that higher mutual conductance is secured and at lower plate voltages than with an equivalent triode. Screen-grid tubes are sometimes operated as space-charge tetrodes by using the screen grid as the control grid.

81. Power Amplifiers Using Pentodes.—The various types of power amplifiers previously discussed all require a fairly large grid-excitation voltage in order to deliver appreciable power outputs. The trend in broadcast receiver design in about 1928 began to eliminate the first stage of audio-frequency amplification and the power amplifier was operated directly from the output of the detector. This was made possible by the change from grid to plate detection, which enabled much larger radio-frequency voltages to be handled, with the consequent increase in the audio-frequency output of the detector. The elimination of the first audio-frequency amplifier reduced the possibilities of hum caused by the lack of sufficient filtration in the *B* supply, and also reduced the amount of frequency distortion. The demand for larger power outputs from the loud-speaker continued to increase and in many cases the maximum output of the power stage was unattainable owing to the inability of the detector to furnish the required grid excitation without serious distortion. Either a power tube capable of furnishing large outputs with a much

lower grid voltage than that required by the available triodes was necessary, or else the first audio stage had to be restored.

The pentode was offered as an inexpensive solution to this problem. This tube avoids the power limitation present in the screen-grid tetrode resulting from secondary emission by the introduction of a third grid between the screen grid and plate, as shown in Fig. 137. This "suppressor grid" is at the same potential as the cathode and serves to repel the secondary electrons back to the plate when the plate potential falls below that of the screen grid. The screen or second grid can now be operated at the same potential as the plate, which increases the mutual conductance of the tube. This grid is similar to the others in construction as it is unnecessary to shield the control grid from the fluctuations in the plate potential at audio frequencies. In pentodes used for radio-frequency amplification this shielding is necessary and the screen-grid construction is then similar to that employed in tetrodes for the same purpose.

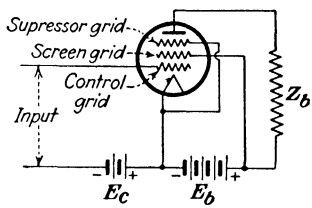


FIG. 137.—Connections of a pentode with a suppressor grid.

The suppressor grid is connected to the mid-point of the filament within the tube, or in the later types using an equipotential cathode, it is brought out to an individual pin in the tube base and the connection is made externally. This construction is more flexible and permits triode operation of the tube by reconnecting the grids, as in the case of the type 46 tube previously discussed. Triple-grid tubes are available which lend themselves to various modes of operation. When used as a pentode, the grid nearest the cathode serves as a control grid, while the second and third are used as a screen grid and suppressor grid, respectively. For Class A triode operation the second and third grids are tied to the plate and the first grid again serves as the control electrode. This connection reduces the internal resistance and amplification factor to values low enough for satisfactory operation as a Class A power amplifier. For Class B operation in push-pull circuits the amplification factor should be high enough to permit the tube to operate with zero bias, as self-bias is impractical. This is accomplished by tying grids one and two together and

using them as the control grid, while the third grid is tied to the plate.

The I_p-E_p characteristics of a typical pentode are shown in Fig. 138. With 250 volts on both the plate and screen grid and a negative bias of -16.5 volts, the amplification factor is 150 and the internal resistance is 60,000 ohms. The proper value of load resistance for this tube is 7000 ohms, which is very much less than r_p. A higher value of load resistance would increase the power output, but the amplitude distortion would also increase. This will be evident if we reduce the slope of the load

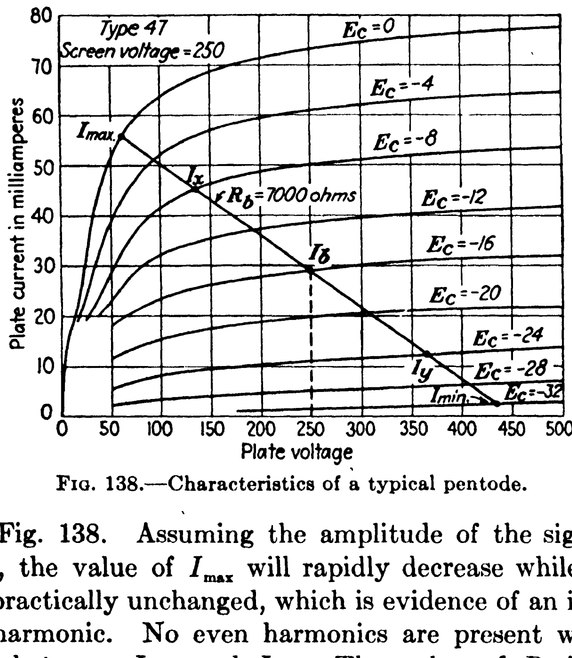


FIG. 138.—Characteristics of a typical pentode.

line in Fig. 138. Assuming the amplitude of the signal to be 16 volts, the value of I_{max} will rapidly decrease while I_{min} will remain practically unchanged, which is evidence of an increasing second harmonic. No even harmonics are present when I_b is midway between I_{max} and I_{min}. The value of R_b is usually chosen so as to give this relation. This condition is different than with triodes where the distortion decreased as R_b was made larger in comparison to r_p. It is therefore important in the case of power pentodes to have the load inductance small in order to prevent the impedance of the load from increasing with the frequency. Loud-speakers to be operated by pentodes should therefore have as small a phase angle as possible. The electrodynamic loud-speaker is better in this respect than the other types. In the case of a loud-speaker having appreciable induc-

tance the increase in load impedance at the higher frequencies will cause considerable distortion in this region. A condenser of 0.02 to 0.03 μf in series with about 15,000 ohms is often shunted across the primary of the output transformer to prevent this rise in load impedance at the higher frequencies. This arrangement is sometimes converted into a tone control by using a larger value of resistance and making it variable. However, this will vary the effective load impedance, which is undesirable.

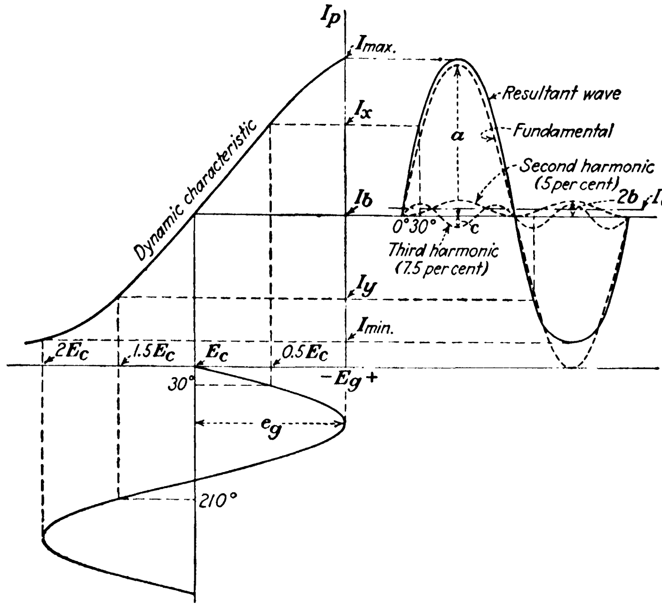


FIG. 139.—Production of harmonics in a pentode when load resistance is too low.

The chief advantage of pentodes as power amplifiers lies in the large power output that can be obtained with a nominal value of signal voltage. The amplitude distortion, mostly third harmonic, is higher than with triodes operating Class A.

82. Graphical Determination of Distortion and Power Output of Pentodes.—The dynamic characteristic of a pentode will have an appearance similar to Fig. 139 when the load resistance has too low a value. This produces both second and third harmonics in the output. The magnitudes of these harmonics can be determined from the I_p-E_p characteristics of Fig. 138. The currents I_x and I_y are the values of plate current obtained when

e_g is $0.5E_c$ and $1.5E_c$, respectively, assuming that the maximum value of e_g is just equal to the biasing voltage E_c. Referring to Fig. 139, the amplitudes of the fundamental, second, and third harmonics are a, b, and c.

From the diagram,

$$I_{\max} = I_b + a + 2b - c \tag{73}$$

$$I_{\min} = I_b - a + 2b + c \tag{74}$$

Adding (73) and (74) and solving for b,

$$b = \frac{I_{\max} + I_{\min} - 2I_b}{4} \tag{75}$$

Subtracting (74) from (73),

$$I_{\max} - I_{\min} = 2a - 2c \tag{76}$$

Also,

$$I_x = I_b + 0.5a + 0.5b + c \tag{77}$$

$$I_y = I_b - 0.5a + 0.5b - c \tag{78}$$

Subtracting (78) from (77),

$$I_x - I_y = a + 2c \tag{79}$$

Adding (76) to (79) and solving for a, we get

$$a = \frac{I_{\max} - I_{\min} + I_x - I_y}{3} \tag{80}$$

Eliminating a from (76) and (79), the amplitude of the third harmonic is found to be

$$c = \frac{2(I_x - I_y) - (I_{\max} - I_{\min})}{6} \tag{81}$$

The percentages of the second and third harmonics in terms of the fundamental are then

$$\%\,2\text{nd} = 100\frac{b}{a} = 75\frac{I_{\max} + I_{\min} - 2I_b}{I_{\max} - I_{\min} + I_x - I_y} \tag{82}$$

$$\%\,3\text{rd} = 100\frac{c}{a} = 100\frac{I_x - I_y - \frac{1}{2}(I_{\max} - I_{\min})}{I_{\max} - I_{\min} + I_x - I_y} \tag{83}$$

The power output at fundamental frequency is

$$P = i_p^2R_b = \left(\frac{a}{\sqrt{2}}\right)^2 R_b = \frac{(I_{\max} - I_{\min} + I_x - I_y)^2}{18}R_b \tag{84}$$

The total harmonic output will be

$$\%\ \text{total} = \sqrt{(\%\ 2\text{nd})^2 + (\%\ 3\text{rd})^2} \tag{85}$$

The above equations assume that harmonics higher than the third may be neglected, which is usually the case under ordinary conditions of operation. The expressions are also applicable to triodes in cases where the operating conditions may produce appreciable amounts of third harmonic.

It will be observed from (82) that the second harmonic will be zero when $I_{max} + I_{min} = 2I_b$, which can be brought about by the proper choice of load resistance. For values of R_b larger than this, $I_{max} + I_{min}$ becomes less than $2I_b$. This results in a negative value for (82), which means that the phase of the second harmonic is reversed from the position shown in Fig. 139. The negative half-wave of i_p is then greater than the positive half wave which causes a *reduction* in the average value of the plate current, and I_b' will now be less than I_b. This will change the operating point on the characteristics if the resistance of the primary winding of the output transformer is appreciable, as discussed on page 182. In this case the position of the load line will have to be lowered.

The effect of the load resistance on the harmonic content and power output of the pentode of Fig. 138 is shown in Fig. 140. As will be observed, the second harmonic disappears at about 7000 ohms where a reversal of its phase takes place. The use of push-pull circuits employing pentodes will therefore produce little improvement from the standpoint of distortion as compared to triodes under similar conditions, since the suppression of even harmonics can be accomplished in the case of pentodes by the proper choice of load impedance. The remaining sizable third harmonic will not be greatly affected by the push-pull circuit as the composite characteristics are similar in shape to those of a single tube. There are, of course, other advantages to the push-pull circuit aside from reduced distortion, such as the reduction of alternating current hum, the absence of degenerative effects at low frequencies when using self-bias, etc.

83. Direct-coupled Amplifiers.—In order to amplify extremely low frequencies, or voltages which vary slowly and which may have periods of no change in value, a direct-coupled amplifier must be used. These are often referred to as direct-current

amplifiers. One arrangement is shown in Fig. 141*a*. Separate *B* batteries must be used for each stage. The grid of the next stage is connected to a point in the battery which is at a slightly negative potential with respect to the filament of this stage,

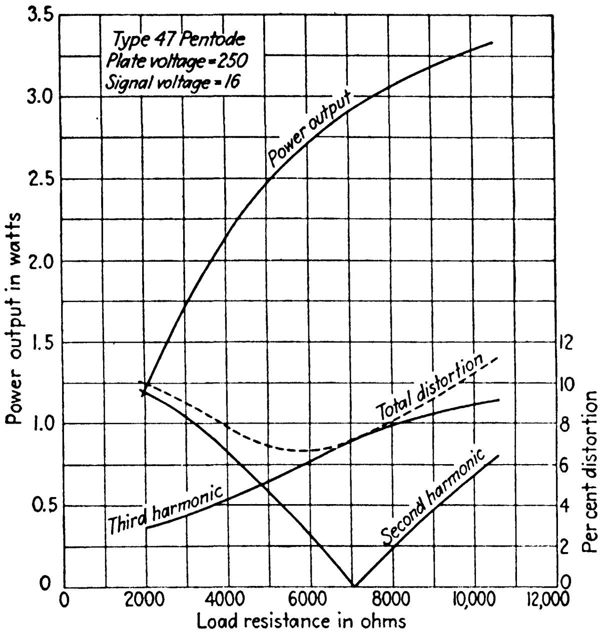


FIG. 140.—Effect of load resistance on distortion and power output in a typical pentode.

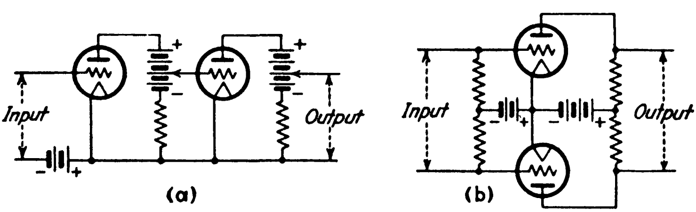


FIG. 141.—Direct-current amplifiers.

giving it the proper negative bias. Amplifiers of this type are rather difficult to keep in balance owing to the voltage drift in the various batteries with use. The balanced amplifier of Fig. 141*b* avoids these difficulties to a considerable extent, as changes in the various battery voltages affect both tubes alike. These circuits are more temperamental and the amplification obtainable

is usually much lower than a conventional alternating-current amplifier of a similar number of stages.[6]

Another form of direct-coupled amplifier which has been used in broadcast receivers is shown in Fig. 142. This arrangement, due to Loftin and White,[7] obtains the various grid and plate voltages from the *IR* drops across portions of a high resistance connected across the source of high voltage. Instability is avoided by obtaining the bias required for the first tube as the difference between the voltage drops across R_1 and R_2. The connection from the cathode through C_1 to an adjustable point

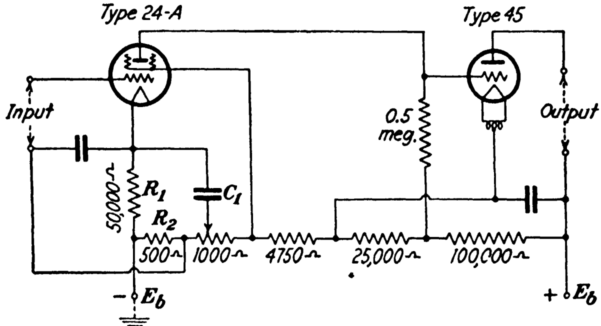


FIG. 142.—Direct-coupled amplifier of the Loftin-White type.

on the 1000-ohm resistance is for the purpose of balancing out any hum that may be present owing to poor filtration of the high-voltage supply. The principle involved in the latter is the same as the bridge circuit of Fig. 88 used in the measurement of μ.

Amplifiers of these types have negligible phase shift and are satisfactory to use with an oscillograph. As the oscillograph vibrator is a current-operated device of relatively low resistance, amplitude distortion due to curvature in the characteristic of the last tube must be guarded against by placing sufficient resistance in series with the plate so as to straighten the characteristic. A tube having as low a value of r_p as possible should be used.

[6] For additional information on amplifiers of this type see J. M. Eglin, A Direct-current Amplifier for Measuring Small Currents, *Jour. Optical Soc. Amer.*, vol. 18, p. 393, May, 1929; also articles by Taylor and Kerr, *Rev. Sci. Inst.*, vol. 4, p. 28, January, 1933, and O. H. A. Schmitt, vol. 4, p. 661, December, 1933.

[7] *Proc. I.R.E.*, vol. 18, p. 669, April, 1930.

The effect of low plate resistance can be secured by using several tubes in parallel in the final stage.

84. Feed-back in Amplifiers.—If a portion of the amplified output energy of an amplifier is fed back into the input circuit in the proper phase so as to reinforce the input voltage, the gain will be greatly increased. This action is termed *regeneration.* If the amount of feed-back is sufficient, a continuous oscillation will be produced. In audio-frequency amplifiers this is termed "singing" or "howling."

Let us assume that a voltage e is impressed upon an amplifier free from phase shift which has an amplification of A. The output will be Ae. If a fraction s of the output is fed back into the input, the voltage fed back will be sAe, and the net output will be $Ae(1 - s)$. The amount thus fed back to the input circuit will again be amplified and a fraction of this will again be returned to the input circuit. This process can be more readily seen from the following:

$$\begin{array}{lll}
e & \rightarrow Ae & \rightarrow Ae(1-s) \\
 & \swarrow & \\
sAe & \rightarrow sA^2e & \rightarrow sA^2e(1-s) \\
 & \swarrow & \\
s^2A^2e & \rightarrow s^2A^3e & \rightarrow s^2A^3e(1-s) \\
\cdots & \cdots & \cdots \\
s^nA^ne & \rightarrow s^nA^{n+1}e & \rightarrow s^nA^{n+1}e(1-s)
\end{array}$$

The total output will be the sum of the terms in the right-hand column which is

$$E_t = Ae(1 - s)(1 + As + A^2s^2 + A^3s^3 + \cdots) \tag{86}$$

The right-hand term of (86) is an infinite series, which will be convergent if $As < 1$. Its sum will then be

$$\frac{1}{1 - As} \tag{87}$$

so that (86) can be written

$$E_t = Ae\frac{1 - s}{1 - As} \tag{88}$$

The *regenerative* or *feed-back* amplification will be

$$A_r = \frac{E_t}{e} = A\frac{1 - s}{1 - As} \tag{89}$$

This expression assumes that the voltage fed back to the input circuit was exactly in phase with e and that the output was reduced by the amount fed back. If the input impedance to the amplifier is infinite so that feed-back represents no reduction of the output, (89) becomes

$$A_r = \frac{A}{1 - As} \tag{90}$$

As the term As in the denominator of either of these two expressions approaches unity, the regenerative amplification becomes enormously large, and is infinite when $As = 1$. Under this condition $E_t/e = \infty$, which means that E_t can be finite, even though e is zero. In other words, an output voltage will be present, even though no external input voltage is applied, which is the condition of a sustained oscillation. Therefore, if instability is to be avoided in an amplifier, As must be kept less than unity. The term A may be looked upon as the "repeater" amplification, or the amount of gain possessed by the amplifier in the absence of feed-back. The higher the gain of an amplifier, the smaller the percentage of feed-back required to cause instability in the form of oscillation, or howling.

This is readily demonstrated in installations of public-address systems where feed-back in the form of sound energy from the loud-speakers can enter the microphone. As the gain control of the amplifier is turned up, more of the amplified output energy enters the input circuit so that a point is reached—assuming sufficient gain on the part of the amplifier—where a continuous howl is produced. Either the gain must be kept below the singing point, or else the feed-back must be reduced by relocating the microphone or loud-speakers, or, as a third possibility, the directional features of the latter must be improved. Serious distortion is present just below the singing point and the gain must be usually kept about 4 db below this point for good quality. A microphonic tube in an early stage of the amplifier may also cause trouble of this sort if exposed to the sound waves or mechanical vibrations from the loud-speaker. This was a fairly common occurrence in the earlier types of radio sets.

The possibilities of acoustic feed-back are present only when the amplified output is converted into sound energy in the immediate vicinity of the amplifier. Electrical sources of feed-

back which inadvertently couple the input and output circuits must also be guarded against, and with increasing care as the overall gain becomes larger. In audio-frequency amplifiers the chief source of such coupling is in the common source of plate-voltage supply. The internal impedance of this source is traversed by the plate currents of all the tubes, so that the IZ drops across this common impedance are introduced into the plate circuit of the first tube, as shown in Fig. 143. These voltages may either aid or oppose the e.m.f. $\mu_1 e_s$ acting in this circuit, since the total voltage will be the vector sum of all the voltages present. The phase relations of these voltages will change with the frequency because of the phase shift caused by the interstage-coupling reactances, so that the regenerative effects may increase

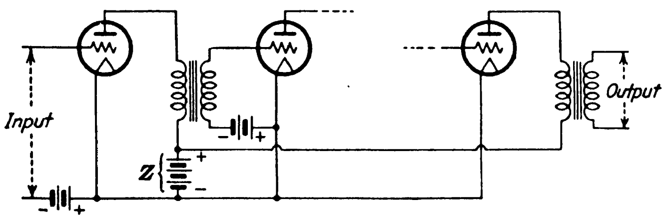


FIG. 143.—Circuit of a multistage amplifier having coupling between the first and last stages due to the use of a common source of plate voltage.

the gain considerably in one range of frequencies and reduce it in another. This can cause serious frequency distortion, even though no instability is produced. When the plate currents of several amplifier tubes all flow through the same common impedance, the voltage drop which the current from one stage develops across the common impedance will transfer energy to all of the other stages, and an accurate solution of the problem is extremely complex. Fortunately, the transfer of energy from the plate circuit of the last tube to that of the first is usually all that need be considered, as the differences in energy levels between any other pairs of tubes is small compared with that between the first and the last.

Computations involving feed-back may be made by means of (89) or (90). Where phase shift must be taken into account, A and s, which are voltage (or current) ratios, will have a phase angle as well as magnitude and must be used as complex numbers.

When the amount of regeneration is sufficient to cause sustained oscillations, the frequency is governed by the resonant

frequency of the circuits involved. In resistance-coupled amplifiers feed-back sufficient to cause instability manifests itself as "motor-boating," a name given to it because of the similarity of the sounds produced in the loud-speaker to the exhaust of a motor boat. Since there is no resonant circuit present, the wave form of the oscillation is badly distorted and the frequency is governed chiefly by the time constant of the grid leak and blocking condenser. The frequency is very low, usually of the order of 1 cycle per second. A disturbance applied to the grid of one tube is amplified and fed back in the proper phase so as to augment the initial impulse, which continues to increase in an

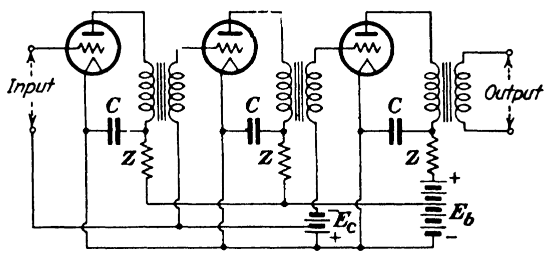


FIG. 144.—Multistage amplifier with filter circuits in each plate lead to minimize effects of common battery coupling.

abrupt fashion until one of the tubes involved reaches cut-off. The amplifying action of this tube then ceases until the blocking potential leaks off of the grid, when the cycle of events is again repeated.

Motor-boating is particularly troublesome with resistance-coupled amplifiers operated from rectified alternating current because of the high internal impedances of such sources at low frequencies. When so operated, these amplifiers must have their individual plate circuits properly filtered.

Feed-back due to self-bias has already been discussed in Sec. 76.

85. Prevention of Feed-back.—Frequency distortion caused by feed-back due to common coupling in the plate-voltage supply can be minimized by shunting a very large condenser across this source of voltage and thereby reducing its internal impedance. At low frequencies the reactance of the condenser becomes large so that it is no longer effective and frequency distortion may be present at these frequencies. The feed-back responsible for this

form of distortion may sometimes prove beneficial and improve the drooping gain in this region if the phase relations are proper.

A more effective method of avoiding the effects of common coupling is to confine the alternating components of plate current to their individual stages and thus prevent them from flowing through the source of plate supply, as illustrated in Fig. 144. Here the plate circuit of each tube is provided with a filter composed of a by-pass condenser C and an impedance Z in series with each plate lead. The condenser provides a path of low impedance to the flow of the alternating-current component of plate current, while the high impedance Z tends to prevent its flow through the common plate battery. The impedance Z is usually

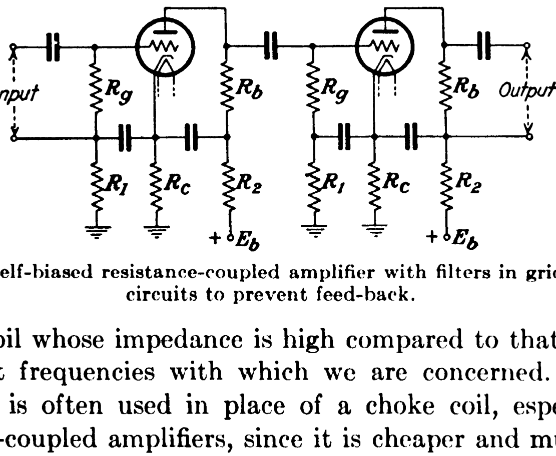


FIG. 145.—Self-biased resistance-coupled amplifier with filters in grid and plate circuits to prevent feed-back.

a choke coil whose impedance is high compared to that of C for the lowest frequencies with which we are concerned. A high resistance is often used in place of a choke coil, especially in resistance-coupled amplifiers, since it is cheaper and much more compact. The drop in voltage across these resistances is usually of no consequence if a rectifier is used as a source of plate power, as a dropping-resistance must be used in any event to secure the reduced voltage required by the first two tubes in Fig. 144.

Feed-back effects when self-bias is used must also be avoided. Figure 145 shows how both grid and plate circuits may be filtered in a self-biased resistance-coupled amplifier using heater-type tubes. The negative bias is secured through the voltage drop in R_c, while R_1 and R_2 serve as filtering resistances. The negative terminal of E_b is connected to ground.

Push-pull circuits operating Class A are free from coupling difficulties resulting from the use of a common source of plate power as the alternating component of plate current does not flow through this source. With Class B and Class AB operation

the plate-supply current fluctuates at double the signal frequency, so that disturbances caused by these fluctuations must be guarded against by the use of adequate filters.

Electromagnetic and electrostatic couplings between circuit elements of the various stages can be taken care of by careful location of the component parts or by individually shielding each stage. The latter is not usually resorted to at audio frequencies unless very high gain is to be employed. The amplifier as a whole is usually shielded so as to prevent it from picking up external electrical disturbances. This is of particular importance when the input is at a very low energy level, as is the case with the output of the various types of high-quality microphones. Inductive disturbances due to near-by lighting circuits may induce voltages in the first stage of an unshielded amplifier, or

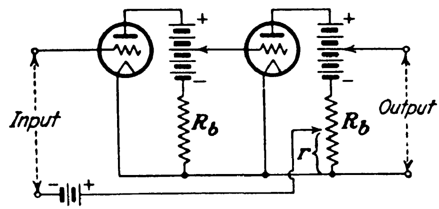


FIG. 146.—Direct-coupled amplifier with adjustable positive feed-back.

in its input leads, which are comparable to, or even greater than the output of the microphone.

Feed-back due to the capacitance between the grid and plate of a triode is negligible at audio frequencies and will be discussed in Chap. VIII.

86. Amplifiers Employing Feed-back.—The use of feed-back or regeneration in the proper phase so as to reinforce the impressed signal will greatly increase the amplification as the value of As in (90) is made to approach unity. This principle, discovered by E. H. Armstrong,[8] has been extensively employed to increase the sensitivity of a triode when used as a detector. The magnitude and phase of the energy fed back are functions of the frequency when reactive elements are involved, so that serious frequency distortion would be produced by its use in most types of audio-frequency amplifiers.

One method of using feed-back at audio frequencies is shown in the direct-coupled amplifier of Fig. 146. The drop across a

[8] *Proc. I.R.E.*, vol. 3, p. 215, 1915.

portion of the output resistance is in series with the input voltage so that the resultant voltage applied to the grid of the first tube will be the sum of these two voltages. The value of r must be critically adjusted to secure high gain. Two tubes are needed to secure the proper phase relations, as introducing a portion of the output of the first tube into its own input circuit by direct coupling would oppose the signal voltage instead of aiding it, as was pointed out when methods of securing self-bias were discussed.

If an amplifier is constructed so as to have considerably more gain than is actually required and the excess gain is then reduced to the desired amount by using negative feed-back, which means that s is made negative and the denominator of (90) becomes greater than unity, some rather remarkable characteristics can be obtained.[9] Amplifiers employing this principle can be made to have constant amplification over a very wide range of frequencies and with an amazing reduction in distortion. Noise produced within the amplifier is greatly reduced and changes in the amplification with variations in the plate voltage are almost negligible. In order to secure these advantages, very careful control is required of the phase shifts within the amplifier and feed-back circuits. These circuits must be designed so that the proper phase shift is obtained for the useful frequency band as well as for a wide range of frequencies above and below it if instability is to be avoided.[10]

With phase shift present, both A and s in (90) are complex quantities. The absolute magnitude of the regenerative amplification is

$$|A_r| = \frac{|A|}{|1 - As|} \tag{91}$$

Defining

$$As = |As|\underline{/\phi} \tag{92}$$

(91) can be written

[9] See papers by H. S. BLACK, Stabilized Feed-back Amplifiers, *Elec. Eng.*, vol. 53, p. 114, January, 1934; and *Bell Lab. Rec.*, vol. 12, p. 290, June, 1934; also E. PETERSON, J. G. KREER, and L. A. WARE, Regeneration Theory and Experiment, *Proc. I.R.E.*, vol. 22, p. 1191, October, 1934.

[10] H. NYQUIST, Regeneration Theory, *Bell System Tech. Jour.*, vol. 11, p. 126, January, 1932.

$$|A_r| = \frac{|A|}{|1 - |As|(\cos\phi + j\sin\phi)|} = \frac{|A|}{\sqrt{(1 - |As|\cos\phi)^2 + |As|^2\sin^2\phi}}$$

$$= \frac{|A|}{\sqrt{1 - 2|As|\cos\phi + |As|^2}} \tag{93}$$

In order to determine the effect of feed-back upon the stability of amplification, suppose that owing to some cause or other the repeater amplification changes by an amount δA, which will then produce a corresponding change in the regenerative amplification of δA_r. Viewing the stability as the ratio of δA_r to A_r, the differential of (93) with respect to A is, assuming small variations,

$$\delta|A_r| = \frac{1 - |As|\cos\phi}{[1 - 2|As|\cos\phi + |As|^2]^{3/2}}\delta A \tag{94}$$

The ratio of $\delta|A_r|$ to $|A_r|$ is

$$\frac{\delta|A_r|}{|A_r|} = \frac{1 - |As|\cos\phi}{1 - 2|As|\cos\phi + |As|^2}\left[\frac{\delta|A|}{|A|}\right] \tag{95}$$

It will be observed from (95) that if $|As|\cos\phi = 1$, the numerator of the right-hand term is zero and the gain stability is perfect, assuming differential variations in $|A|$. In a similar

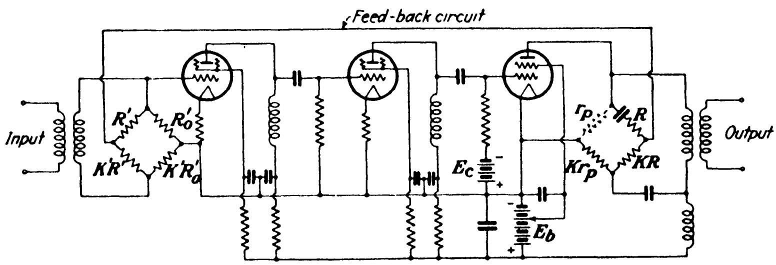


FIG. 147.—Circuit of a negative feed-back amplifier with the feed-back connected through bridges so as to prevent undesired reactions with input and output circuits.

manner it is possible to investigate the stability with regard to small variations in s and ϕ.

The circuit of a negative feed-back amplifier is shown in Fig. 147, the input and output being coupled by means of balanced bridges. This prevents the circuits connected to the input and

output from affecting the amount of feed-back. The magnitude and phase of the feed-back can be controlled by the adjustment of the bridge. The variation of the gain with frequency for the amplifier of Fig. 147, taken from Black's paper, is shown in Fig. 148. The upper curve shows the characteristic of the amplifier with the reversed feed-back removed. The effect of feed-back in reducing the frequency distortion is strikingly shown by

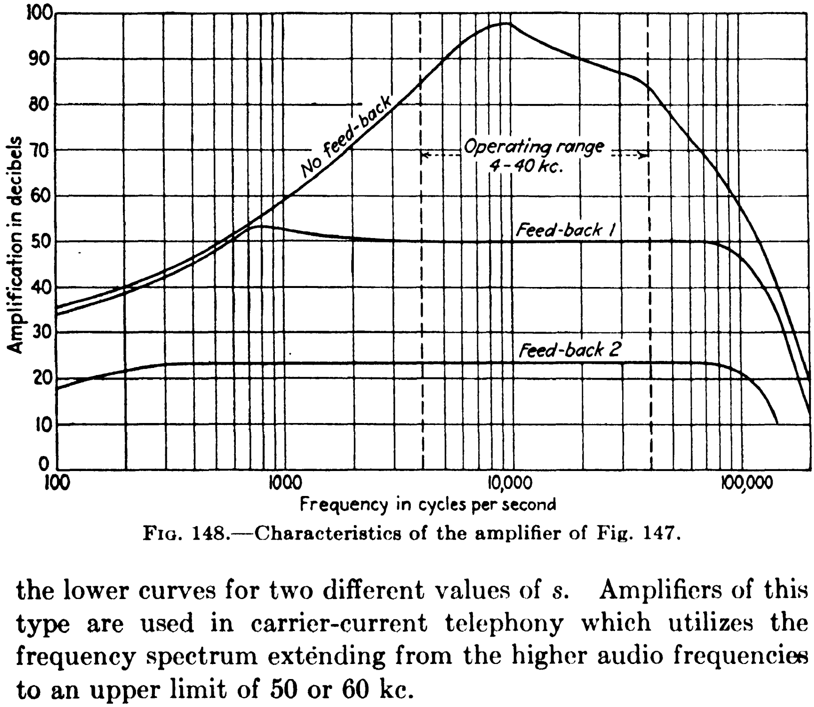


FIG. 148.—Characteristics of the amplifier of Fig. 147.

the lower curves for two different values of s. Amplifiers of this type are used in carrier-current telephony which utilizes the frequency spectrum extending from the higher audio frequencies to an upper limit of 50 or 60 kc.

By placing suitable equalizers in the feed-back circuit the gain-frequency characteristics of the amplifier can be made to compensate for the loss-frequency characteristics of the telephone line in which the amplifier is placed.

87. Telephone Repeaters.—The amplifiers previously considered are capable of transmitting and amplifying signals in one direction only. Where two-way transmission is desired over a single pair of wires, the input and output terminals of the amplifier must be electrically isolated so as to prevent singing. This is accomplished by making the input and output terminals of the amplifier two pairs of diagonally opposite points of a bridge

circuit, as shown in Fig. 149*a*. The output terminals of the amplifier are connected across points *ab* while the input terminals are connected across points *cd*. If the bridge is balanced so that $Z_a/Z_b = Z_1/Z_2$, none of the output voltage of the amplifier can appear across the input terminals. But if a voltage *e* is introduced into either Z_1 or Z_2, by someone speaking into the apparatus at

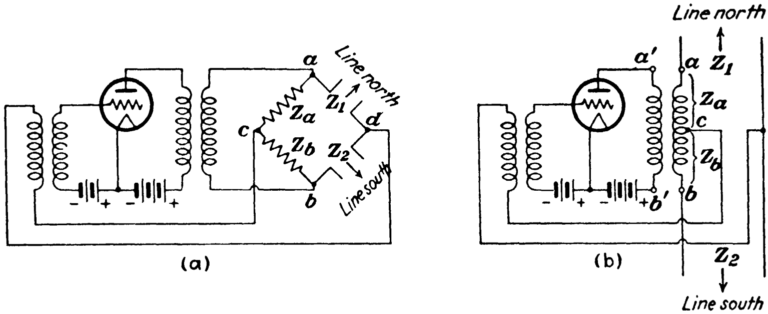


FIG. 149.—(*a*) Amplifier and bridge circuit connected to give two-way transmission. (*b*) 21-type repeater circuit evolved from (*a*).

the end of the line, the bridge is no longer balanced for a voltage inserted in either of these arms and a portion of this voltage is impressed across the input terminals *cd*. The amplified output will appear across *ab* and hence is impressed across the two lines in series. If $Z_1 = Z_2$, the ratio arms Z_a and Z_b can be replaced by the two halves of the output-transformer secondary, as shown in Fig. 149*b*. This is known as a "21-type"-repeater circuit. In telephone practice it is very desirable to keep the impedance of each conductor the same and balanced with respect to ground. Inserting Z_a and Z_b in one side of the line would upset this condition, so that one-half of the coils Z_a and Z_b are removed and placed in the other side of the line, as shown in Fig. 150. The coils Z_a and Z_b will both have the same number of turns. This type transformer is often called a *hybrid transformer*. The windings are all on a common core with the secondary coils connected so that 1-2-3-4 is the series-aiding connection.

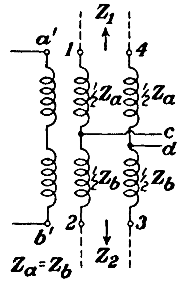


FIG. 150.—Connections of a hybrid transformer used with telephone repeaters.

The 21-type repeater requires that the impedances Z_1 and Z_2 of the two lines shall be the same. Any deviation from this

relationship will introduce positive feed-back, so that the gain A must be reduced in order that As may remain less than unity to avoid singing. This restriction would prevent the use of a repeater between lines whose impedance-frequency characteristics were dissimilar. This limitation may be avoided by the use of a 22-type-repeater circuit illustrated in Fig. 151. The circuit consists of two 21-type repeaters, except that each line is now balanced by means of a network having the same impedance-frequency characteristics. Either line can be unbalanced without the production of singing, but not both. This type of repeater is the most commonly used, owing to its greater flexibility. For a more detailed discussion of these types of repeaters the reader

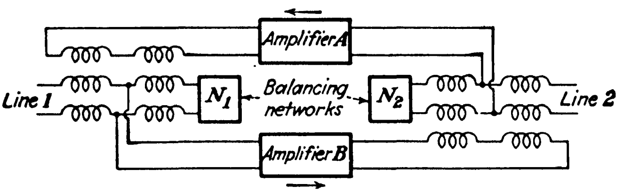


FIG. 151.—22-type-repeater circuit.

is referred to "Transmission Circuits for Telephonic Communication," by K. S. Johnson, Chap. XIV.

The arrangement of Fig. 151 is used to terminate an ordinary telephone line for two-way radio telephony. Assuming the telephone subscriber to be talking over line 1, the output of amplifier B would be supplied to the radio transmitter. The radio receiver would be connected to the input of amplifier A. The right-hand hybrid transformer with its associated line and balancing network would, of course, be absent. Difficulties caused by feed-back will be present if the output of the transmitter is picked up by the receiving set. This is avoided by the use of two different frequencies for transmitting and receiving, and often by geographical separation of the transmitting and receiving stations as well.

Problems

1. A resistance-coupled amplifier using type 40 triodes has the following circuit constants per stage: $R_b = R_c = 0.5$ megohm, $C = 0.1$ μf, $E_b = 180$ volts, $E_c = -1.5$ volts. The tube constants are $\mu = 30$, $r_p = 150{,}000$ ohms, $C_{gp} = 8.0$ $\mu\mu$f, $C_{gf} = 2.8$ $\mu\mu$f. The distributed capacitance shunted across C_{gf} due to wiring, etc., is 14.2 $\mu\mu$f. What is the

voltage amplification per stage at frequencies of 10, 100, 1000, and 10,000 cycles?

TABLE I

$E_c = 0$		$E_c = -1$		$E_c = -2$	
E_p	I_p^*	E_p	I_p^*	E_p	I_p^*
20	0.57	60	0.21	130	0.08
40	0.85	80	0.39	150	0.23
60	1.20	100	0.61	170	0.47
80	1.58	120	0.90	200	0.90
100	1.96	140	1.22	250	1.70

* Milliamperes.

2. The characteristics of one triode unit of a type 79 duplex tube are given in Table I. This tube is used in a resistance-coupled amplifier of the following constants: $R_b = R_c = 0.25$ megohm, $E_b = 250$ volts, $E_c = -1$ volt. The impressed signal has a maximum value of 1 volt and the frequency is such as to make the reactance of the blocking condenser and the input capacitance of the following tube negligible. What is the output voltage of the tube across R_c?

3. The two triode units of the tube in Problem 2 are paralleled by connecting the grids, plates, and cathodes together. What will be the output voltage, assuming all the other conditions to be the same as before?

4. A triode having $\mu = 30$ and $r_p = 60{,}000$ ohms has a pure inductance L_b in its plate circuit. What must be the value of L_b so that the voltage amplification at 100 cycles is 90 per cent of the value at 1000 cycles?

5. In Fig. 106 the impedance Z_1 is a pure resistance of 100,000 ohms. The impedance Z_2 has an alternating-current resistance of 1000 ohms and an inductance of 50 henrys. The condenser Z_c is a pure capacitance of the proper size to resonate with Z_2 at 60 cycles. The constants of the tube are $\mu = 10$, $r_p = 10{,}000$ ohms. What is e_2/e_g at 60 cycles?

6. An audio-frequency transformer has a primary winding of 4000 turns and a secondary winding of 12,000 turns. The net core area is 0.9 sq. in. and the mean length of the core is 6 in. The transformer is connected to a type 56 triode biased so that $\mu = 14$, $r_p = 9000$ ohms, $I_b = 0.003$ amp. If the flux density in the transformer core is 30,000 lines per square inch, what is the open-circuit voltage across the secondary if 1 volt at 60 cycles is applied to the grid of the tube? At 1000 cycles? Neglect the primary resistance and assume the incremental permeability of the core material to be $\frac{1}{10}$ the direct-current permeability.

7. A high-quality, 3:1 ratio audio-frequency transformer has a primary inductance of 120 henrys and a secondary inductance of 1080 henrys. The equivalent leakage inductance referred to the primary is 0.25 henry. The remainder of constants, using the symbols of Fig. 111, are as follows: $r_1 = 1960$ ohms, $r_2 = 9450$ ohms, $C_1 = 310$ $\mu\mu$f, $C_2 = 80$ $\mu\mu$f, $C_m = 30$ $\mu\mu$f, $L_c = 120$ henrys, $R_c = 12{,}000$ ohms. The transformer is connected to a

tube whose constants are $\mu = 14$ and $r_p = 9000$ ohms. If the value of C_0 connected across the transformer secondary is 100 $\mu\mu$f, at what frequencies will resonant peaks occur in the amplification curve, assuming C_m (*a*), to aid the magnetic couplings and (*b*), to oppose it?

8. In Problem 7, what will be the voltage amplification at these resonant peaks under each condition? Compare these with value that would be obtained with an ideal 3:1 transformer.

9. A two-stage, transformer-coupled amplifier has a magnetic type of loud-speaker inserted directly in the plate circuit of the last tube. The impedance of the loud-speaker is $10{,}000\underline{/30°}$. The input transformer to the first tube has a step-up ratio of 3:1 and the coupling transformer between tubes is 2.5:1. The latter has a primary impedance of $40{,}000\underline{/90°}$. If the constants of both tubes are $\mu = 8$ and $r_p = 10{,}000$, ohms, what will be the voltage across the loud-speaker if 0.05 volt is impressed across the primary of the input transformer? What is the gain of the amplifier in decibels?

10. A source of e.m.f. has an internal impedance of $6 + j8$ ohms. Permanently shunted across this source is a pure resistance of 20 ohms. What value of load impedance will absorb the maximum amount of power from

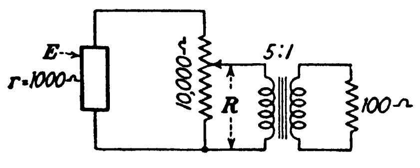


FIG. *A*.

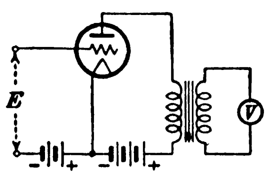


FIG. *B*.

this source if the ratio of R to X of the load is constant and is given by $Z = R + j0.5R$?

11. A radio set in a hotel has a dynamic loud-speaker fed from a power tube whose constants are $\mu = 3.8$ and $r_p = 2000$ ohms. This tube is connected to the loud-speaker by means of a step-down transformer of ratio 16:1, which may be considered as ideal. The loud-speaker impedance is $8 + j6$ ohms. The management wishes to operate 5 magnetic-type loud-speakers located in various dining rooms from this set. It is proposed to connect these five speakers in parallel and shunt them across the primary of the output transformer in the set, using leads of negligible impedance. It is desired that these five speakers shall make as much noise as possible. The dynamic loud-speaker is to remain connected to the set for monitoring purposes and the resultant reduction of its output is immaterial. The magnetic speakers can be wound for any impedance, the phase angle remaining constant at 60 degrees. What impedance should be specified for the magnetic-type speakers?

12. A source of e.m.f. has an internal resistance of 20 ohms. Connected to this source is a transformer of the following constants: $Z_p = 5 + j200$, $Z_s = 50 + j2000$, $Z_m = 0 + j600$. What is the vector expression for a load impedance Z_L to be connected across Z_s so that the load will absorb the maximum possible amount of power?

13. The internal resistance of a source of e.m.f. E in Fig. A is 1000 ohms. Shunted across this source is a 10,000-ohm potentiometer. Find the value of R in order that the power in the 100-ohm resistance shall be a maximum. Assume the 5:1 transformer to be ideal.

14. A low-reading alternating-current voltmeter has a resistance of 100 ohms. It is desired to increase the sensitivity of the voltmeter by the use of an amplifier as shown in Fig. B so that the voltage E to be measured will produce larger meter deflections than would be obtained by connecting the meter directly across the source of E. The output transformer may be considered ideal and is chosen to have the best possible ratio. Show that unless the tube used for this purpose has a ratio of $\mu/\sqrt{r_p}$ greater than 0.2 the voltage amplification will be less than unity.

TABLE II

$E_c = 0$		$E_c = -20$		$E_c = -30$		$E_c = -40$		$E_c = -60$	
E_p	I_p^*	E_p	I_p^*	E_p	I_p^*	E_p	I_p^*	E_p	I_p^*
25	12	100	7	140	7	170	5	230	2
50	37	120	21	160	21	190	14	250	7
75	71	140	42	180	42	210	30	270	15
100	111	160	70	200	69	230	53	290	28
125	155	180	102	220	100	250	82	310	50
150	204	200	136	240	132	270	113	330	78
175	258	225	184	260	168	290	145	350	110

$E_c = -80$		$E_c = -90$		$E_c = -100$		$E_c = -120$	
E_p	I_p^*	E_p	I_p^*	E_p	I_p^*	E_p	I_p^*
310	2.5	340	2	380	1.5	450	1.5
330	8	360	6	400	6	460	3
350	18	380	13	420	13	475	6
370	32	400	24	440	24	500	15
390	51	420	40	460	42	525	27
410	76	440	63				

* Milliamperes.

15. The characteristics of a type 2A3 triode are given in Table II. The tube is to be operated with a plate supply of 250 volts and a grid bias of -40 volts. Plot curves of the power output and second harmonic for the following values of load resistance in ohms: 1000, 2000, 3000, 4000, 5000. The maximum value of the impressed signal is 40 volts and the output transformer has a 1:1 ratio. Neglect the third harmonic and the ir drop in the tube filament.

16. Determine μ and r_p from the characteristic curves of the above tube for $E_b = 250$ and $E_c = -40$ and compute the power output from equation (59). Check these results with those obtained in Problem 15.

17. Two type 2A3 triodes are operated push-pull with a plate supply of 300 volts and a grid bias of -60 volts. From the data given in Table II construct the necessary composite characteristics to determine the power output and per cent third harmonic for the following values of plate-to-plate load resistance, assuming the output transformer ratio is 1:1 (total primary to secondary): 2000, 3000, 4000, and 5000 ohms. The maximum value of the signal impressed on each grid is 60 volts.

18. *a.* A triode having $\mu = 4$ and $r_p = 2000$ has a magnetic-type loud-speaker connected directly in its plate circuit. The loud-speaker impedance is $2000 + j1000$ at 100 cycles. If a signal of 10 volts r.m.s. is applied to the grid find the power absorbed by the loud-speaker and the voltage across it.

b. If the *C* battery used with this tube is then replaced by a biasing resistance of $R_c = 1000$ ohms, as in Fig. 130 but with the by-pass condenser omitted, what will be the power and voltage across the loud-speaker?

19. The biasing resistance in Problem 18 is shunted with a 4-μf condenser. Find the power and voltage across the loud-speaker under this condition.

20. The various items of the circuit shown in Fig. *C* are as follows: $\mu = 5$, $r_p = 2000$ ohms, $R_c = 500$ ohms, $R_2 = 7200$ ohms. What transformation ratio N_s/N_p will result in maximum power in R_2? If $e_s = 20$ volts, what will be the value of the maximum power?

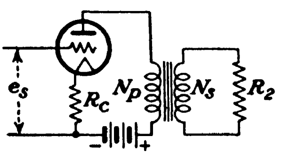


Fig. *C*.

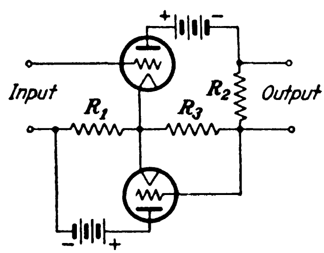


Fig. *D*.

21. The characteristics of a type 89 triple-grid tube are given in Table III when connected for pentode operation with 250 volts applied to the screen grid. The tube is operated with a plate supply of 250 volts and a grid bias of -25 volts. Determine the power output, percentage second and third harmonics, and total percentage of harmonics for a load resistance coupled to the tube by means of a 1:1 output transformer, ranging from 3000 to 10,000 ohms in steps of 1000 ohms. Plot the values obtained against load resistance, assuming a signal of 25 volts, maximum. What seems to be the most desirable value of load resistance?

22. Two type 89 pentodes are operated push-pull with the same voltages as in Problem 21. From the data given in Table III construct the necessary composite characteristics to determine the power output and percentage third harmonic for plate-to-plate load resistances coupled to the two tubes by means of a 1:1 output transformer, ranging from 4000 to 20,000 ohms

in steps of 4000 ohms. Plot the values obtained against load resistance. The maximum value of the signal impressed on each grid is 25 volts. What seems to be the most desirable value of load resistance? Compare the results with those of Problem 21 for a single tube.

TABLE III

$E_c = 0$		$E_c = -12.5$		$E_c = -25$		$E_c = -37.5$		$E_c = -50$	
E_p	I_p^*	E_p	I_p^*	E_p	I_p^*	E_p	I_p^*	E_p	I_p^*
25	51.5	25	32.9	25	18.8	50	11.3	50	4.3
35	58.7	50	42.6	50	24.3	100	12.8	100	5.0
50	65.6	75	47.0	75	27.2	150	13.6	150	5.2
75	71.7	100	49.0	100	28.3	200	14.0	200	5.4
100	75.4	150	52.0	175	30.5	300	15.0	300	5.7
150	79.5	200	54.1	250	32.2	400	15.9	400	6.0
200	81.8	300	57.9	325	33.8	500	16.6	500	6.2

* Milliamperes.

23. Repeat Problem 15, taking the third harmonic into account.

24. Find the voltage amplification of the circuit in Fig. 146 assuming both tubes to be alike and $\mu = 10$, $r_p = 10{,}000$ ohms, $R_b = 90{,}000$ ohms, $r = 1000$ ohms.

25. Derive an expression for the voltage amplification of the circuit in Fig. *D*, assuming both tubes to be alike.

CHAPTER VIII

INPUT IMPEDANCE OF A TRIODE

88. Equivalent Circuit of a Triode.—The equivalent circuit of Fig. 87 neglects the interelectrode capacitances of the tube, which become of increasing importance at higher frequencies. Even at audio frequencies these capacitances result in an input impedance which cannot be neglected. The equivalent circuit of a triode when these capacitances are included is shown in Fig. 152. The grid is assumed to be biased negatively by a sufficient amount so that the grid current i_g that flows is not due to electrons attracted to the grid. The latter case can be taken into account by considering a resistance r_g to be shunted across C_{gf}. Actually, all of these capacitances should have resistance associated with them in order to take dielectric losses and leakage into account. However, under most conditions of operation their effects are small in comparison to those of the capacitances so that they will be neglected in the following treatment.

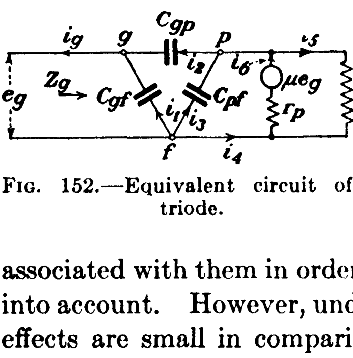


FIG. 152.—Equivalent circuit of a triode.

The circuit of Fig. 152 is unique in that two e.m.fs. are acting which are related in such a way that the voltage of one is entirely controlled by the voltage of the other.

89. Input Impedance of the Grid Circuit.—The impedance looking into the grid-filament terminals of the tube will be given by the vector ratio of the applied voltage e_g to the resultant current i_g. It will be assumed that e_g is small so that μ and r_p may be regarded as constants. The expression for Z_g can be obtained by applying Kirchhoff's laws to the equivalent circuit and solving for i_g in terms of the impressed voltage and the circuit constants. The assumed positive directions through the circuit are shown by the arrows in Fig. 152.

The following equations are obtained:

$$\dot{i}_6 = \dot{i}_4 + \dot{i}_5 \tag{1}$$
$$\mu e_g = \dot{i}_6 r_p + \dot{i}_5 Z_b \tag{2}$$
$$\dot{i}_5 Z_b = -\frac{\dot{i}_3}{j\omega C_{pf}} \tag{3}$$
$$\dot{i}_2 = \dot{i}_3 + \dot{i}_4 \tag{4}$$
$$\frac{\dot{i}_1}{j\omega C_{gf}} = \frac{\dot{i}_2}{j\omega C_{gp}} + \frac{\dot{i}_3}{j\omega C_{pf}} \tag{5}$$
$$\dot{i}_g = \dot{i}_1 + \dot{i}_2 \tag{6}$$
$$e_g = \frac{\dot{i}_1}{j\omega C_{gf}} \tag{7}$$

Substituting (1) in (2),

$$\mu e_g = \dot{i}_4 r_p + \dot{i}_5(r_p + Z_b) \tag{8}$$

Substituting the value of $\dot{i}_5$ from (3) in (8),

$$\mu e_g = \dot{i}_4 r_p - \frac{\dot{i}_3}{j\omega C_{pf} Z_b}(r_p + Z_b) \tag{9}$$

From (4), $\dot{i}_4 = \dot{i}_2 - \dot{i}_3$, and substituting this in (9), we get

$$\mu e_g = \dot{i}_2 r_p - \dot{i}_3\left(r_p + \frac{r_p + Z_b}{j\omega C_{pf} Z_b}\right) \tag{10}$$

From (5),

$$\dot{i}_3 = C_{pf}\left(\frac{\dot{i}_1}{C_{gf}} - \frac{\dot{i}_2}{C_{gp}}\right)$$

and substituting this value for $\dot{i}_3$ in (10) gives us

$$\mu e_g = \dot{i}_2 r_p - C_{pf}\left(\frac{\dot{i}_1}{C_{gf}} - \frac{\dot{i}_2}{C_{gp}}\right)\left(r_p + \frac{r_p + Z_b}{j\omega C_{pf} Z_b}\right) \tag{11}$$

Substituting the value of $\dot{i}_2$ obtained from (6) in (11),

$$\mu e_g = (\dot{i}_g - \dot{i}_1) r_p - C_{pf}\left(\frac{\dot{i}_1}{C_{gf}} - \frac{\dot{i}_g - \dot{i}_1}{C_{gp}}\right)\left(r_p + \frac{r_p + Z_b}{j\omega C_{pf} Z_b}\right) \tag{12}$$

Substituting the value of $\dot{i}_1$ obtained from (7) in (12),

$$\mu e_g = (\dot{i}_g - j\omega C_{gf} e_g) r_p - C_{pf}\left(j\omega e_g - \frac{\dot{i}_g - j\omega C_{gf} e_g}{C_{gp}}\right)\left(r_p + \frac{r_p + Z_b}{j\omega C_{pf} Z_b}\right)$$

$$= i_g r_p - j\omega C_{gf} r_p e_g + \left[\frac{r_p C_{pf} i_g - j\omega C_{gf} C_{pf} r_p e_g}{C_{gp}} - j\omega C_{pf} r_p e_g + \frac{(r_p + Z_b)(C_{pf} i_g - j\omega C_{gf} C_{pf} e_g)}{j\omega C_{gp} C_{pf} Z_b} - \frac{(r_p + Z_b) e_g}{Z_b} \right] \quad (13)$$

Collecting all terms containing e_g on the left and those containing i_g on the right, we get

$$e_g\left(\mu + j\omega C_{gf} r_p + j\omega \frac{C_{gf} C_{pf} r_p}{C_{gp}} + j\omega C_{pf} r_p + C_{gf} \frac{r_p + Z_b}{C_{gp} Z_b} + \frac{r_p + Z_b}{Z_b}\right) = i_g\left(r_p + \frac{r_p C_{pf}}{C_{gp}} + \frac{r_p + Z_b}{j\omega C_{gp} Z_b}\right) \quad (14)$$

Since $Z_g = e_g/i_g$, the input impedance is

$$Z_g = \frac{r_p + \dfrac{r_p C_{pf}}{C_{gp}} + \dfrac{r_p + Z_b}{j\omega C_{gp} Z_b}}{\left(\mu + C_{gf} \dfrac{r_p + Z_b}{C_{gp} Z_b} + \dfrac{r_p + Z_b}{Z_b}\right) + j\omega\left(C_{gf} r_p + \dfrac{C_{gf} C_{pf} r_p}{C_{gp}} + C_{pf} r_p\right)}$$

$$= \frac{r_p(C_{gp} + C_{pf}) - j\dfrac{1}{\omega}\left(\dfrac{r_p}{Z_b} + 1\right)}{\mu C_{gp} + (C_{gf} + C_{gp})\left(\dfrac{r_p}{Z_b} + 1\right) + j\omega r_p(C_{gf} C_{gp} + C_{gp} C_{pf} + C_{pf} C_{gf})} \quad (15)$$

Equation (15) is the vector expression for the input impedance in terms of the tube constants and the impedance of the load Z_b, and assumes that the grid bias is sufficiently negative so as to prevent the flow of electrons to the grid. Three cases arise, depending upon the nature of the load.

90. Case 1. $Z_b = R_b$.—When the load is a pure resistance R_b, (15) may be written

$$Z_g = \frac{a - j\dfrac{b}{\omega}}{c + j\omega d} = \frac{ac - bd}{c^2 + \omega^2 d^2} - j\frac{1}{\omega\dfrac{c^2 + \omega^2 d^2}{bc + \omega^2 ad}} \equiv R_g - j\frac{1}{\omega C_g} \quad (16)$$

where the coefficients have the values

$$\left.\begin{aligned} a &= r_p(C_{gp} + C_{pf}) \\ b &= \frac{r_p}{R_b} + 1 \\ c &= \mu C_{gp} + (C_{gf} + C_{gp})\left(\frac{r_p}{R_b} + 1\right) \\ d &= r_p(C_{gf} C_{gp} + C_{gp} C_{pf} + C_{pf} C_{gf}) \end{aligned}\right\} \quad (17)$$

In order to gain an idea as to the relative magnitudes of these coefficients, let us assume a circuit having the following constants:

$$\begin{aligned} \mu &= 10 \\ r_p &= 10{,}000 \\ R_b &= 100{,}000 \\ C_{gf} &= C_{gp} = C_{pf} = 10\ \mu\mu\text{f} \\ \omega &= 10^6. \end{aligned}$$

The foregoing tube capacitances are somewhat larger than those of the ordinary receiving tube. From (17)

$$\begin{aligned} a &= 2 \times 10^{-7} \\ b &= 1.1 \\ c &= 1.22 \times 10^{-10} \\ d &= 3 \times 10^{-18} \end{aligned}$$

and from (16)

$$R_g = \frac{(2.44 - 0.33) \times 10^{-17}}{(1.4884 + 0.0009) \times 10^{-20}} = 1417 \text{ ohms}$$

$$C_g = \frac{(1.22^2 + 0.0009) \times 10^{-20}}{(1.342 + 0.006) \times 10^{-10}} = 110.5\ \mu\mu\text{f}$$

It is evident that for frequencies such that $\omega < 10^6$, a can be neglected in comparison to b, and d is negligible in comparison to c, so that R_g and C_g may be considered to be independent of the frequency and (16) may be written

$$R_g = \frac{ac - bd}{c^2} \tag{18}$$

$$C_g = \frac{c}{b} \tag{19}$$

Substituting the values of c and b in (19),

$$\begin{aligned} C_g &= \frac{\mu C_{gp} + (C_{gf} + C_{gp})\left(\dfrac{r_p}{R_b} + 1\right)}{\dfrac{r_p}{R_b} + 1} \\ &= C_{gf} + C_{gp}\left(1 + \frac{\mu R_b}{r_p + R_b}\right) \end{aligned} \tag{20}$$

The last term in the parenthesis in (20) will be recognized as the expression for the voltage amplification of a resistance-

coupled amplifier. As R_b is made very large compared to r_p, the voltage amplification approaches μ as a limiting value, so that the maximum possible value of grid input capacitance will be

$$C_{g_{max}} = C_{gf} + C_{gp}(1 + \mu) \tag{21}$$

If R_b is zero, R_g is also zero (neglecting dielectric losses) and $C_g = C_{gf} + C_{gp}$. Under this condition the short circuit between the plate and filament places C_{gf} and C_{gp} in parallel and also short-circuits r_p. If C_{gf} and C_{gp} have dielectric losses associated with them due to the imperfect dielectric of the tube base and glass stem, R_g is not zero but will be equal to the dielectric resistance. At low frequencies the dielectric loss causes R_g to be much larger than the values given by (16) and (18).

The term $ac - bd$ appears in the numerator of the expression for R_g. This term becomes zero when $R_b = 0$ and is positive for all other values of R_b, so that the input resistance of a triode will always be positive for a resistance load in the plate circuit.

91. Case 2. $Z_b = R_b + j\omega L_b$.—If the impedance in the plate circuit contains inductive reactance, (15) becomes

$$Z_g = \frac{C_{gp} + C_{pf} - j\frac{1}{\omega}\left(\frac{1}{R_b + j\omega L_b} + \frac{1}{r_p}\right)}{\frac{\mu C_{gp}}{r_p} + (C_{gf} + C_{gp})\left(\frac{1}{R_b + j\omega L_b} + \frac{1}{r_p}\right) + j\omega(C_{gf}C_{gp} + C_{gp}C_{pf} + C_{pf}C_{gf})}$$

$$= \frac{(R_b + j\omega L_b)(C_{gp} + C_{pf}) - j\frac{1}{\omega} - j\frac{1}{\omega r_p}(R_b + j\omega L_b)}{\frac{\mu C_{gp}}{r_p}(R_b + j\omega L_b) + C_{gf} + C_{gp} + \frac{1}{r_p}(C_{gf} + C_{gp})(R_b + j\omega L_b) + j\omega(R_b + j\omega L_b)(C_{gf}C_{gp} + C_{gp}C_{pf} + C_{pf}C_{gf})}$$

$$= \frac{R_b(C_{gp} + C_{pf}) + \frac{L_b}{r_p} - j\left[\frac{1}{\omega}\left(1 + \frac{R_b}{r_p}\right) - \omega L_b(C_{gp} + C_{pf})\right]}{\frac{R_b}{r_p}(\mu C_{gp} + C_{gp} + C_{gf}) + C_{gf} + C_{gp} - \omega^2 L_b \times (C_{gf}C_{gp} + C_{gp}C_{pf} + C_{pf}C_{gf}) + j\omega\left[R_b(C_{gf}C_{gp} + C_{gp}C_{pf} + C_{pf}C_{gf}) + \frac{L_b}{r_p}(\mu C_{gp} + C_{gp} + C_{gf})\right]} \tag{22}$$

This may be written

$$Z_g = \frac{a - jb}{c + jd} = \frac{ac - bd}{c^2 + d^2} - j\frac{bc + ad}{c^2 + d^2} \equiv R_g - j\frac{1}{\omega C_g} \tag{23}$$

where

$$C_g = \frac{c^2 + d^2}{\omega(bc + ad)} \tag{24}$$

The coefficients have the values

$$\left.\begin{aligned} a &= R_b(C_{gp} + C_{pf}) + \frac{L_b}{r_p} \\ b &= \frac{1}{\omega}\left(1 + \frac{R_b}{r_p}\right) - \omega L_b(C_{gp} + C_{pf}) \\ c &= \frac{R_b}{r_p}(\mu C_{gp} + C_{gp} + C_{gf}) + C_{gf} + C_{gp} - \omega^2 L_b \times \\ &\qquad (C_{gf}C_{gp} + C_{gp}C_{pf} + C_{pf}C_{gf}) \\ d &= \omega\left[R_b(C_{gf}C_{gp} + C_{gp}C_{pf} + C_{pf}C_{gf}) + \right. \\ &\qquad \left. \frac{L_b}{r_p}(\mu C_{gp} + C_{gp} + C_{gf})\right] \end{aligned}\right\} \tag{25}$$

The numerator of the expression for R_g is $ac - bd$, which from (25) is equal to

$$N = R_b C_{gp}^2 + \frac{R_b^2 + \omega^2 L_b^2}{r_p}(\mu C_{gp}^2 + C_{gp}^2 + \mu C_{gp}C_{pf}) - \frac{L_b}{r_p}\mu C_{gp} \tag{26}$$

The value of N will be negative, and hence R_g will be negative when

$$\mu L_b > r_p R_b C_{gp} + (R_b^2 + \omega^2 L_b^2)(\mu C_{gp} + C_{gp} + \mu C_{pf}) \tag{27}$$

which is a quadratic equation in L_b. There will in general be two values of L_b that will make R_g equal to zero, and for values between these two, R_g will be negative. If R_b is large the solution for L_b at a particular frequency may be imaginary, in which case there is no value of L_b which will make R_g zero or negative.

The variations in R_g with L_b are shown in Fig. 153 for a typical triode, using several values of R_b. The input capacitance for the same tube and circuit conditions is shown in Fig. 154. Both curves are for $\omega = 5 \times 10^6$.

Regeneration occurs when the impedance of the plate circuit results in a negative value for R_g, and output energy is fed into the input circuit through C_{gp}. If R_g is sufficiently negative to

annul all of the positive resistance in the source of e_g, oscillations will occur. This is responsible for the production of undesired oscillations in tuned radio-frequency amplifiers using triodes.

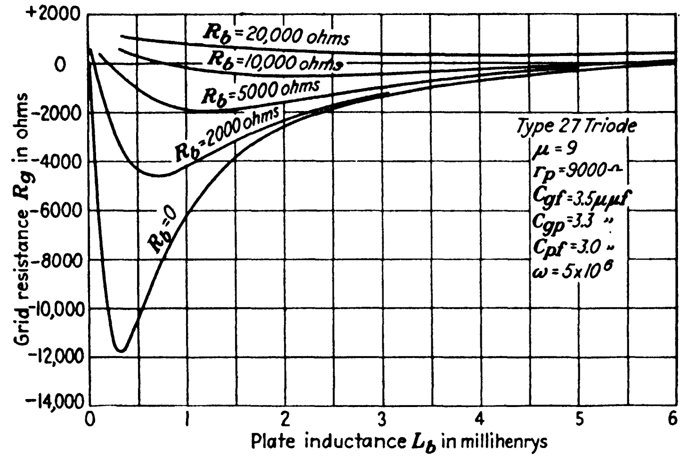


FIG. 153.—Curves showing the variation of input resistance for a typical triode as the plate-circuit inductance is varied.

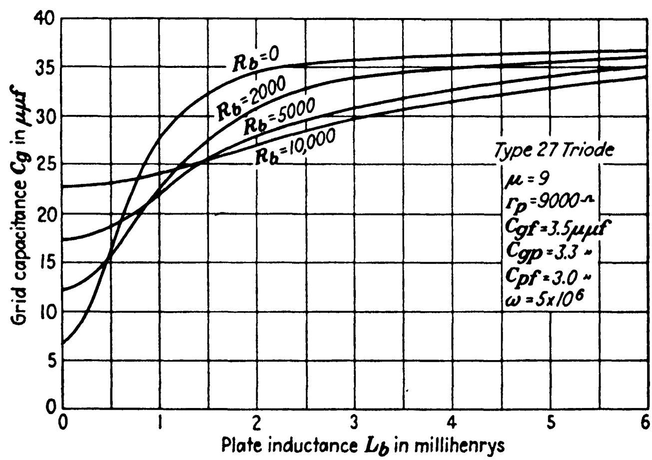


FIG. 154.—Curves showing the variation of input capacitance for a typical triode as the plate-circuit inductance is varied.

Screen-grid tubes reduce the effective value of C_{gp} to a point where the usual values of inductance in the plate circuit are not sufficient to cause oscillations. In tubes of this type the input capacitance is substantially constant and equal to the C_{gf} plus the capacitance between the control and screen grids.

The introduction of negative resistance into the input circuit by means of a variable inductance in series with the plate has been extensively used in one type of regenerative detector. This circuit employs a variometer in the plate circuit which causes R_g to vary in the manner shown in Fig. 153. Regeneration control is accomplished either by increasing L_b from a low value, or else by initially setting L_b at its maximum value and then reducing it. The signal strength greatly increases as the circuit approaches oscillation. At the present time regenerative circuits are chiefly used for short-wave reception where their ability to oscillate is a decided advantage. At broadcast frequencies high sensitivity can be more conveniently obtained and controlled by the use of radio-frequency amplification preceding the detector.

92. Case 3. $Z_b = R_b - jX_b$.—When the impedance in the plate circuit contains capacitive reactance, the expression for Z_g can be obtained by substituting $R_b - jX_b$ in (15) for Z_b. The expression is again of the form

$$Z_g = \frac{a - jb}{c + jd} = \frac{ac - bd}{c^2 + d^2} - j\frac{bc + ad}{c^2 + d^2} \equiv R_g - j\frac{1}{\omega C_g} \tag{28}$$

where

$$\left.\begin{aligned} a &= R_b(C_{gp} + C_{pf}) - \frac{X_b}{\omega r_p} \\ b &= \frac{1}{\omega}\left(1 + \frac{R_b}{r_p}\right) + X_b(C_{gp} + C_{pf}) \\ c &= \frac{R_b}{r_p}(\mu C_{gp} + C_{gp} + C_{gf}) + C_{gf} + C_{gp} + \\ &\qquad \omega X_b\,(C_{gf}C_{gp} + C_{gp}C_{pf} + C_{pf}C_{gf}) \\ d &= \omega R_b(C_{gf}C_{gp} + C_{gp}C_{pf} + C_{pf}C_{gf}) - \\ &\qquad \frac{X_b}{r_p}(\mu C_{gp} + C_{gp} + C_{gf}) \end{aligned}\right\} \tag{29}$$

These equations are similar to (25), except that all the signs preceding X_b are now reversed. The terms in the numerator of the expression for R_g will now be positive, so that if the impedance Z_b in the plate circuit contains condensive reactance R_g will always be positive and regeneration in the positive sense cannot occur. This positive value of R_g may be sufficient to reduce appreciably the amplifier gain, especially at radio frequencies.

Problems

1. The triode unit of a 2A6 tube has the following constants: $\mu = 100$, $r_p = 91{,}000$ ohms, $C_{gf} = C_{gp} = 1.7$ $\mu\mu$f, $C_{pf} = 3.8$ $\mu\mu$f. Plot a curve showing the variation of input capacitance with plate-load resistance for values of R_b from zero to 500,000 ohms.

2. The constants of a type 56 triode are $\mu = 13.8$, $r_p = 9500$ ohms, $C_{gf} = C_{gp} = 3.2$ $\mu\mu$f, $C_{pf} = 2.2$ $\mu\mu$f. If the external resistance R_b in series with the plate is 2000 ohms, for what range of values of inductance L_b in series with the plate will the input resistance of the tube be negative, if $\omega = 10^7$?

3. In Problem 2, what is the minimum value for R_b in order that the input resistance shall not be negative for any value of L_b?

CHAPTER IX

RADIO-FREQUENCY AMPLIFIERS FOR RECEPTION

93. Types of Amplifiers.—Radio-frequency amplifiers for reception are operated as Class A devices. They may be subdivided into tuned or untuned amplifiers, depending on the width of the frequency band to which they respond. An untuned amplifier is one that gives fairly uniform response over a wide range of frequencies. A tuned amplifier is one whose constants are adjusted so that a narrow band of frequencies is amplified and all other frequencies lying outside of this selected band are discriminated against. Both of these types may be further classified as to the nature of the coupling means employed.

94. Untuned Amplifiers.—The various types of circuits discussed in the chapter on audio-frequency amplifiers can be used for radio-frequency amplification, but the difficulties due to tube-input capacitance and distributed capacitance of the coils make it much harder to secure uniform amplification over a very wide range of frequencies. If resistance coupling is attempted, the input capacitance of the following tube approaches a short circuit across the plate load at moderate values of radio frequency so that this type of coupling can be used only for low values of radio frequency.

Impedance-coupled amplifiers are more satisfactory but are limited as to the width of the frequency band over which uniformly high amplification can be secured. Input and distributed capacitance tune the coil to parallel resonance so that uniform amplification is obtained only over the range of frequencies for which the parallel resonant impedance of the plate load is large compared to r_p of the tube. Figure 155 shows the voltage amplification for a triode with a choke coil in its plate circuit. The upper curve shows the effects of the capacitance shunted across the coil, which in this case was 22 $\mu\mu$f. The lower curve shows the amplification that would be obtained if this capacitance were absent.

Transformer-coupled amplifiers can also be used, but it is difficult to secure tight coupling between the primary and secondary without the use of an iron core, so that the leakage inductances of the two windings will be relatively large. The limitations of iron cores have already been discussed in Chap. III. Resonant effects are pronounced and somewhat complicated in transformers due to electrostatic as well as electromagnetic coupling which exists between the two windings.

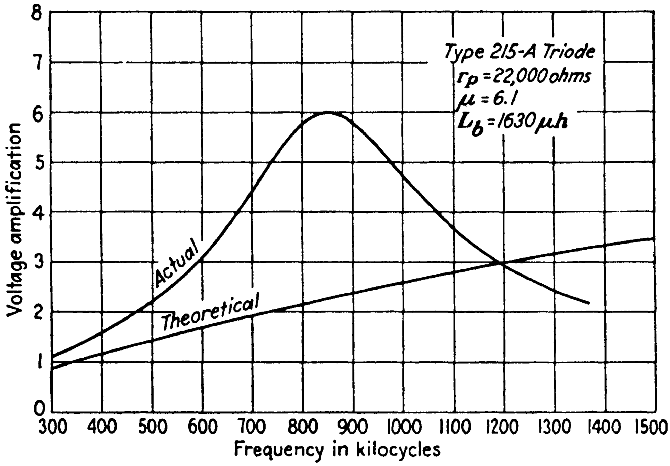


FIG. 155.—Variation of amplification with frequency in a choke-coupled amplifier.

The transformation ratio of an untuned transformer on open circuit is given by

$$a = \frac{M}{L_p} = k\sqrt{\frac{L_s}{L_p}} \tag{1}$$

where M = mutual inductance.
L_p = primary inductance.
L_s = secondary inductance.
k = coefficient of coupling.

This expression assumes that the distributed capacitances of the two windings are negligible and that there is no capacitive coupling between them, assumptions that are only valid at low frequencies.

Untuned amplifiers find their chief applications in various forms of laboratory apparatus such as vacuum-tube oscillators of various types. The radio-frequency amplifiers used in receiving

sets must be capable of discriminating against unwanted signals, so that tuned circuits must be employed to obtain the desired degree of selectivity. Untuned amplifiers are seldom used in modern receiving sets for this reason.

95. Tuned Amplifiers.—The vacuum tube is a source of e.m.f. having relatively high internal resistance and in order to secure

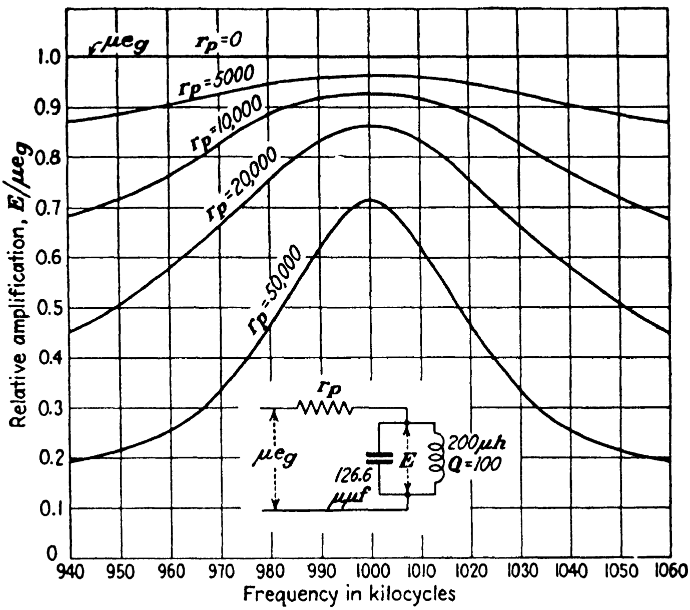


FIG. 156.—Effect of internal resistance of tube on voltage across parallel-resonant circuit in the plate circuit.

the desired selectivity by means of tuned circuits the plate load should be of the nature of a parallel-resonant circuit. Series resonance as a means for obtaining selectivity is effective only when a low resistance source is used. The effect of the internal resistance of the source on the voltage across a parallel-resonant circuit is shown in Fig. 156. If the internal resistance of the tube were zero the voltage output would be constant and equal to μe_g for all frequencies. The impedance of the parallel circuit diminishes rapidly on each side of resonance, and with a tube having a high value of r_p, the voltage across the output circuit falls off abruptly for frequencies above and below resonance. The maximum voltage amplification possible with this type of circuit approaches the value of μ for the tube. Tuned impedance

coupling is sometimes used with screen-grid tubes. These tubes have very high values of μ and r_p and high amplification per stage can be obtained.

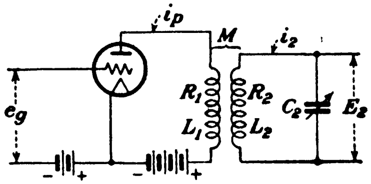


FIG. 157.—Typical circuit of a tuned radio-frequency transformer.

The limiting value of amplification of μ per stage results in insufficient gain with ordinary triodes so that the circuit of Fig. 157 is the type usually used. The voltage across the primary is stepped up by the effective ratio of transformation of the tuned transformer. This ratio of transformation is different from the ratio of turns on the two windings.

The current in the secondary is given by (15), Chap. IV, and is

$$i_2 = \frac{-EZ_m}{(Z_1 + Z_p)(Z_2 + Z_s) - Z_m^2} \tag{2}$$

where

$$\begin{aligned} E &= \mu e_g \\ Z_m &= j\omega M \\ Z_1 + Z_p &= r_p + R_1 + j\omega L_1 \\ Z_2 + Z_s &= R_2 + j\left(\omega L_2 - \frac{1}{\omega C_2}\right) \end{aligned}$$

so that

$$i_2 = \frac{-j\mu e_g \omega M}{(r_p + R_1)R_2 + \frac{L_1}{C_2} - \omega^2(L_1L_2 - M^2) + j\left[(r_p + R_2)\left(\omega L_2 - \frac{1}{\omega C_2}\right) + \omega L_1 R_2\right]} \tag{3}$$

This expression neglects the effects of capacitive coupling between the primary and secondary. If this is appreciable it must be taken into consideration.[1]

The secondary current i_2 will be a maximum, if C_2 is the variable, when the j term in the denominator is zero, which occurs when

$$\frac{1}{\omega C_2} = \omega\left(L_2 + L_1 \frac{R_2}{r_p + R_1}\right) \tag{4}$$

[1] H. DIAMOND and E. Z. STOWELL, Note on Radio-frequency Transformer Theory, *Proc. I.R.E.*, vol. 16, p. 1194, September, 1928.

and is the condition for resonance. The second term in the parenthesis may be regarded as the value of primary inductance reflected into the secondary. Substituting the value of (4) in (3), we get

$$\dot{i}_{2\text{max}} = \frac{-j\mu e_g \omega M}{(r_p + R_1)R_2 + \omega^2\left(M^2 + L_1^2\frac{R_2}{r_p + R_1}\right)} \tag{5}$$

The secondary voltage E_2 at resonance is

$$E_2 = \frac{\dot{i}_{2\text{max}}}{\omega C_2} \tag{6}$$

It will be assumed that this is the maximum value of secondary voltage that can be obtained by adjusting C_2. Actually, E_2 is not a maximum when i_2 is a maximum, but the difference is too small to be taken into account.

The values of R_1 and L_1 are usually small compared with r_p of the tube and can be neglected, resulting in the following approximate expressions which are sufficiently accurate for most purposes:

$$\dot{i}_2 = \frac{-j\mu e_g \omega M}{r_p R_2 + \omega^2 M^2 + j r_p\left(\omega L_2 - \frac{1}{\omega C_2}\right)} \tag{7}$$

At resonance

$$\omega L_2 = \frac{1}{\omega C_2} \tag{8}$$

and the current at resonance is

$$\dot{i}_{2\text{max}} = \frac{-j\mu e_g \omega M}{r_p R_2 + \omega^2 M^2} \tag{9}$$

The voltage amplification is

$$A_v = \frac{E_2}{e_g} = \frac{\dot{i}_{2\text{max}}\frac{1}{\omega C_2}}{e_g} = \frac{\mu\frac{M}{C_2}}{r_p R_2 + \omega^2 M^2} \tag{10}$$

From (8), the voltage amplification may also be written as

$$A_v = \frac{\mu\omega M}{r_p R_2 + \omega^2 M^2}\omega L_2 \tag{11}$$

Differentiating (11) with respect to M, we find that the optimum value of mutual impedance is

$$\omega M = \sqrt{r_p R_2} \tag{12}$$

Substituting (12) in (11), the optimum amplification is

$$A_{\text{opt}} = \frac{\mu \omega L_2}{2\sqrt{r_p R_2}} \tag{13}$$

which is the maximum amplification it is possible to obtain with a given tube and coil.

When M is adjusted to its optimum value, it will be noted that the figure of merit of the tube is $\mu/\sqrt{r_p}$. Therefore, if two tubes have equal values of mutual conductance, the one having the highest amplification factor will give the greatest gain. Tetrodes and pentodes will accordingly produce a greater gain than the ordinary triode. With M less than optimum the gain becomes more nearly proportional to the mutual conductance of the tube. When optimum coupling is employed, the amplification is directly proportional to the ratio of the coil reactance to the square root of its resistance, instead of Q for the coil. With values of M considerably less than optimum, the amplification becomes more nearly proportional to the figure of merit Q of the coil.

From (16), Chap. IV, the impedance looking into the primary coil at resonance is

$$Z = R_1 + j\omega L_1 + \frac{\omega^2 M^2}{R_2} \tag{14}$$

At optimum coupling, neglecting the impedance of the primary coil,

$$Z = r_p \tag{15}$$

or the impedance of the load is equal to the internal resistance of the tube. This condition differs from the impedance-coupled amplifier in that in the latter optimum amplification is approached by making the impedance of the load very large compared with r_p of the tube.

When screen-grid tetrodes and pentodes are used, which sometimes have internal resistances approaching a megohm in value, the coupling used in the transformer is far below the optimum

value, so that $\omega^2 M^2$ in the denominator of (11) can usually be neglected, and (11) becomes approximately

$$A_v = \frac{\mu\omega M}{r_p R_2}\omega L_2 = g_m Q_2 \omega M \tag{16}$$

These tubes enable values of amplification per stage to be obtained which are several times greater than can be obtained with the ordinary triode. This advantage, together with their freedom from oscillation without the use of neutralizing circuits, has caused the use of triodes to be virtually abandoned in the field of radio-frequency amplifiers for receiving sets. Neutralizing circuits will be discussed in Sec. 101. The value of C_{gp} in these tetrodes and pentodes is usually below 0.01 $\mu\mu$f. The connection from the control grid is brought out to a small metal cap on the top of the glass bulb, which is of further assistance in keeping C_{gp} small, as capacitance between the grid and plate leads in the tube base and socket is thereby avoided. A similar construction is used in the case of all-metal tubes. In spite of the small value of C_{gp}, screen-grid tubes may oscillate if the load in the plate circuit is sufficiently inductive. This tendency increases at the higher values of radio frequency, so that amplifiers used for short-wave reception must be designed for a lower gain per stage, if instability is to be avoided, than is the case at lower frequencies. At broadcast frequencies the maximum permissible gain per stage is governed by stability considerations so that the coupling that can be employed with these tubes is usually much lower than the optimum value given by (12).

96. Effect of Mutual Inductance.—The change in the voltage amplification at resonance as the mutual inductance is varied is shown in Fig. 158 for three values of frequency. The curves have rather dull maxima so that the conditions for optimum amplification given by (12) are not very critical. A fixed value of M of about 60 μh would have produced nearly optimum amplification with the tube and circuit constants used for the entire range of frequencies shown.

The sensitivity of a radio set is primarily governed by the number of stages and the gain per stage of the radio-frequency amplifier used. In addition to sensitivity, the selectivity is often of even greater importance. This property is defined as the ability to discriminate against undesired signals differing in

frequency from the desired signal. In sets designed for entertainment purposes, the fidelity is also of importance. Broadcast

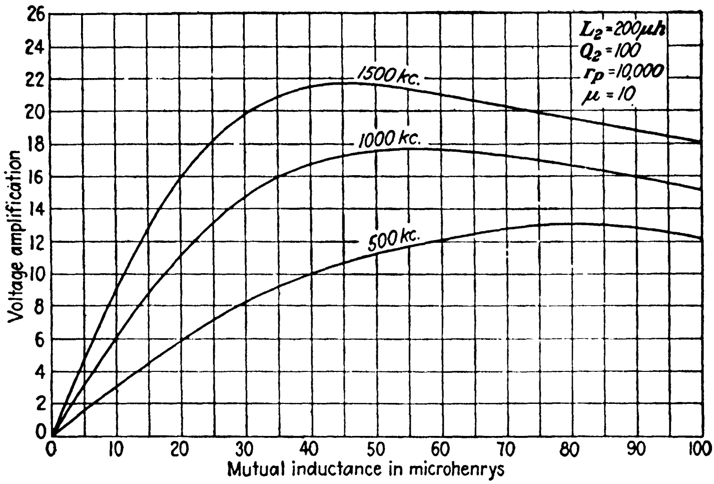


FIG. 158.—Variation of the resonant amplification with mutual inductance.

entertainment requires a band of frequencies approximately 10 kc in width, and to avoid distortion in the radio-frequency

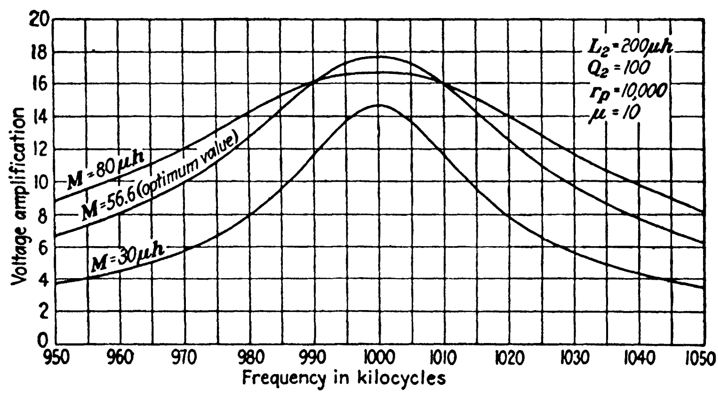


FIG. 159.—Resonance curves of a tuned radio-frequency amplifier for three values of mutual inductance.

amplifier the gain should be approximately constant within this band. The fidelity of sound reproduction is also affected by the characteristics of the detector, audio-frequency amplifier, and loud-speaker.

Figure 159 shows the effect of the amount of mutual inductance upon the resonance curves of a single stage of the type shown in

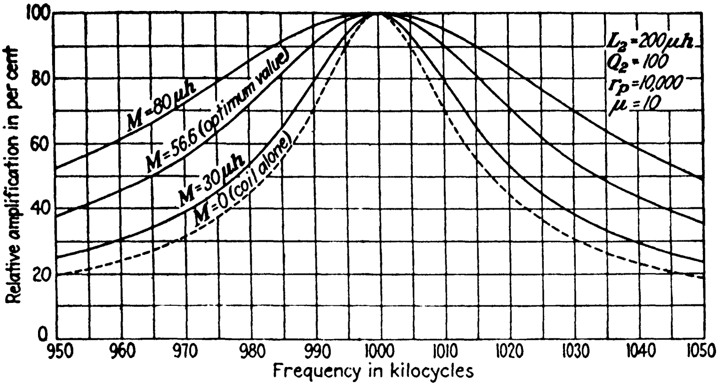


FIG. 160.—Effect of mutual inductance on selectivity.

Fig. 157. A better comparison of the selectivity may be had if the ordinates of the curves of Fig. 159 are expressed as a percentage of the amplification at resonance, as in Fig. 160. As M is

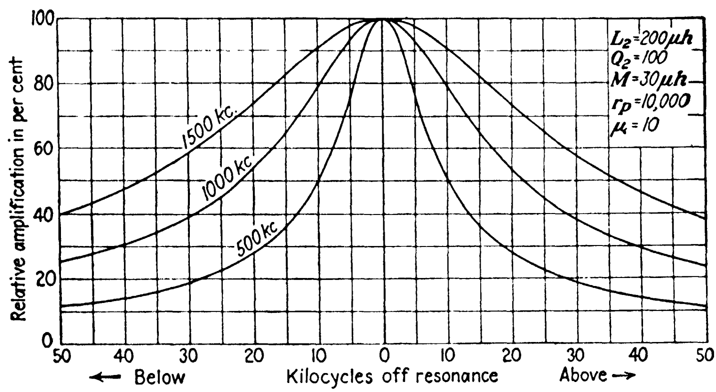


FIG. 161.—Variation in selectivity with frequency.

reduced, the selectivity improves, approaching the resonance curve of the secondary as a limiting value. Greater selectivity would have been obtained if a higher value of Q had been employed in the secondary coil, but with some sacrifice in fidelity. The ideal response curve would be rectangular in shape with a width of about 10 kc. In order to secure fairly uniform response for all frequencies lying 5 kc above and below resonance, the

response curve for the ordinary type of tuned circuit would have to be made entirely too broad for good selectivity. Consequently, the design of an amplifier of this type represents a compromise between these two conflicting requirements. Inadequate selectivity manifests itself to the user very much more readily than impaired fidelity, so that receiving sets using tuned radio-frequency amplification employ a value of mutual inductance between primary and secondary coils which is considerably below the optimum value.

A further conflict between selectivity and fidelity appears when the performance at various frequencies is investigated, as shown in Fig. 161. The value of M remained constant at 30 μh for all three curves. The selectivity at high frequencies is rather poor, but if the circuit constants are changed so as to bring about an improvement in this region, the selectivity at the low frequencies will tend to become too sharp for good fidelity. The use of a fixed value of mutual inductance below the optimum value also causes a considerable reduction in sensitivity at the lower frequencies as may be seen in Fig. 158 along the ordinate $M = 30$. One method of avoiding some of these difficulties would be to vary M automatically with the tuning adjustment, but the resulting mechanical and electrical complications have prevented its use.

The radio-frequency amplifiers in receiving sets designed primarily for the reception of telegraphic code signals are not concerned with fidelity considerations.

97. Combinations of Inductive and Capacitive Coupling.—In order to secure better performance in tuned amplifiers without resorting to moving parts other than the tuning condensers, combinations of inductive and capacitive coupling between stages have been used.[2] By a proper choice of circuit elements it is possible to make the effective coupling vary with the frequency in a predetermined manner. In this way the variation of gain with frequency can be given almost any desired characteristic. The variation in selectivity is also made somewhat more uniform over the tuning range of the amplifier.

[2] For a comprehensive discussion of various types of these circuits the reader is referred to a paper by H. A. Wheeler and W. A. McDonald, Theory and Operation of Tuned Radio-frequency Coupling Systems, *Proc. I.R.E.*, vol. 19, p. 738, May, 1931

Two examples of circuits involving these principles are shown in Fig. 162. In (*a*) the coil L_b has a large value of inductance so that its distributed capacitance C_1, augmented by C_{pf} of the tube, wiring, etc., resonates it to a frequency somewhat below the tuning range of the set. The amplified output current of the tube divides between L_b and the path through the coupling condenser C_m. At low frequencies a larger portion of the output current flows through this second path because of the high impedance offered by L_b as parallel resonance is approached in the latter. The circuit C_mL_1 can be made to approach series resonance in a fashion similar to the double impedance-coupled amplifier of Fig. 106. The apparent inductance looking into L_1 is made up almost entirely of reflected reactance from the tuned secondary

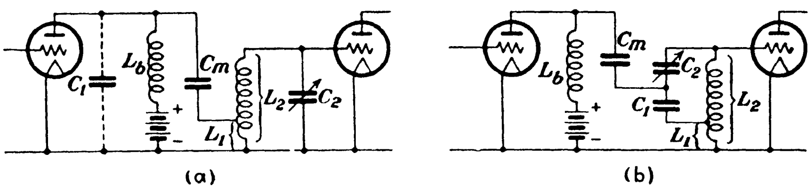


FIG. 162.—Tuned amplifiers using combinations of inductive and capacitive coupling.

circuit L_2C_2. This reflected reactance will vary with the adjustment of the tuning condenser C_2 in a manner similar to Fig. 35 of Chap. II, if the frequency in this figure is replaced by the capacitance of C_2 as the independent variable.

The circuit of Fig. 162*b* achieves the same results in a somewhat different manner. The coil L_b and condenser C_m merely serve as choke coil and blocking condenser of an amplifier using parallel feed. The amplified output current divides between C_1 and C_2 and then recombines to flow through the primary L_1 of the autotransformer. The capacitance of the tuning condenser C_2 is increased as the signal frequency is lowered, which causes a progressive increase in the effective coupling. The condenser C_1 is about twenty times larger than the maximum value of C_2, while L_1 includes about a turn or two of the coil L_2. If C_m is purposely made small the total reactance in the plate circuit will be capacitive. As was shown in Chap. VIII, the resistance term R_g of the tube-input impedance will always be positive for a capacitive plate load, so that reducing the size of C_m affords a

possible means of securing stability without resorting to neutralizing circuits. This is, of course, unnecessary when screen-grid tubes are used, as positive feed-back is here prevented by the reduction of C_{gp} to a value too small to cause oscillations.

The advantages of the circuits just described can be approached by means of the circuit of Fig. 157 if the primary inductance is increased to a large value so that L_1, in conjunction with its distributed capacitance and the output capacitance of the tube used, has a resonant frequency somewhat below the tuning range of the circuit. This becomes a case of two coupled circuits wherein the frequency of the primary is much lower than the impressed frequency. When a conventional primary of few turns is used, its natural frequency is above the highest received frequency. When the secondary is tuned to the lower frequencies, this corresponds to operation at point A in Fig. 62 of Chap. IV. As C_2 is varied, we move along line $a - a'$ and the tuning is quite sharp. When a higher frequency is received, the point of operation shifts along the hyperbola toward point B. The secondary tuning is much broader in this region along the line $a - a'$, and becomes more so as point C is approached at still higher values of received frequency.

But with a large primary the scene of operations is shifted to the other branch of the hyperbola in the third quadrant. Here the secondary tuning would be sharpest at high frequencies and would tend to become broader at the lower frequencies. This would merely transpose the region of broad tuning from one end of the tuning dial to the other if it were not for the fact that the secondary resistance is greatest at high frequencies. Consequently, the greatest resistance is present where the tuning due to coupled-circuit phenomena would tend to be the sharpest, and becomes progressively lower as the coupled-circuit tuning tends to become broader. The effect of resistance predominates and the tuning is sharpest at the low-frequency end, but the variations in selectivity and amplification over the tuning range are much less than with the conventional primary of few turns.

A small fixed condenser is often shunted across L_1, which lowers the natural frequency of the primary and enables a coil of fewer turns to be used. An approach to the behavior of the circuit of Fig. 162*a* may be had by arranging for sufficient capacitive coupling between the primary and secondary coils. This

capacitance between the two windings then takes the place of C_m in the above mentioned diagram. This form of construction has been extensively used in tuned radio-frequency receiving sets employing screen-grid tubes.

The use of a large primary operated above its resonant frequency results in a plate-load impedance which has condensive reactance. When ordinary triodes are used, the grid-input conductance is positive instead of negative, which means that the energy fed back through C_{gp} opposes the impressed signal and weakens it. No troubles with instability will be experienced, but the amplification may be reduced to only a fraction of its theoretical value. Neutralizing circuits must therefore be used to oppose this negative feed-back. This is rather unique in that these circuits are ordinarily used to prevent oscillations. Screen-grid tubes are free from these troubles.

98. Amplifiers Using Tuned Coupled Circuits.—In order to secure a response curve which more nearly approaches the ideal rectangular curve, both the primary and secondary of the coupling

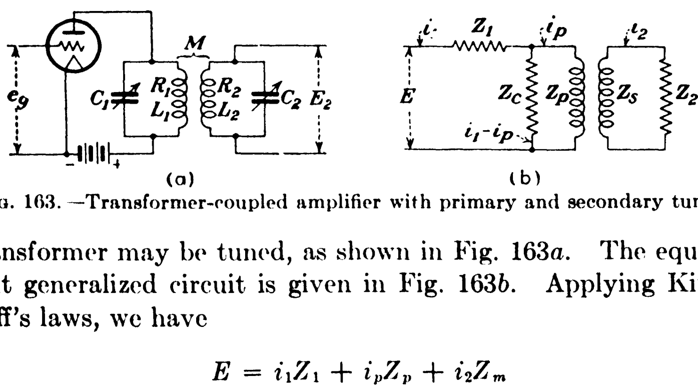


FIG. 163.—Transformer-coupled amplifier with primary and secondary tuned.

transformer may be tuned, as shown in Fig. 163*a*. The equivalent generalized circuit is given in Fig. 163*b*. Applying Kirchhoff's laws, we have

$$E = i_1 Z_1 + i_p Z_p + i_2 Z_m \tag{17}$$

$$E = i_1 Z_1 + (i_1 - i_p) Z_c \tag{18}$$

$$0 = i_2(Z_2 + Z_s) + i_p Z_m \tag{19}$$

From (18)

$$i_1 = \frac{E + i_p Z_c}{Z_1 + Z_c}$$

and substituting this value for i_1 in (17) gives us

$$i_p = \frac{EZ_c - Z_m(Z_1 + Z_c)i_2}{Z_1 Z_c + Z_p Z_1 + Z_p Z_c} \tag{20}$$

Substituting (20) in (19) and solving for i_2, we get

$$i_2 = \frac{-EZ_cZ_m}{(Z_1Z_c + Z_pZ_1 + Z_pZ_c)(Z_2 + Z_s) - Z_m^2(Z_1 + Z_c)}$$
$$= \frac{-EZ_m}{\left[Z_1 + Z_p\left(1 + \frac{Z_1}{Z_c}\right)\right](Z_2 + Z_s) - Z_m^2\left(1 + \frac{Z_1}{Z_c}\right)} \quad (21)$$

Applying this expression to Fig. 163*a*, we have

$$E = \mu e_g \qquad Z_c = \frac{1}{j\omega C_1}$$
$$Z_m = j\omega M \qquad Z_p = R_1 1 + j\omega L_1$$
$$Z_1 = r_p \qquad Z_2 + Z_s = R_2 + j\left(\omega L_2 - \frac{1}{\omega C_2}\right)$$

Substituting these values in (21), the voltage amplification of the circuit is given by

$$A_v = \frac{i_2\frac{1}{\omega C_2}}{e_g} = \frac{\mu\frac{M}{C_2}}{\sqrt{a^2 + b^2}} \quad (22)$$

where

$$a = R_2[R_1 + r_p(1 - \omega^2 L_1C_1)] - \omega(L_1 + r_pR_1C_1)\left(\omega L_2 - \frac{1}{\omega C_2}\right) + \omega^2M^2$$

$$b = \omega R_2(L_1 + r_pR_1C_1) + [R_1 + r_p(1 - \omega^2L_1C_1)]\left(\omega L_2 - \frac{1}{\omega C_2}\right) + \omega^3M^2C_1r_p$$

If both primary and secondary are tuned to resonance the variation in the amplification with the impressed frequency will be similar in appearance to the coupled circuits of Fig. 61. When M is adjusted to its optimum value, the curve has a single maximum, and for greater values of M the response curve will have two peaks; one above and the other below the resonant frequency of the primary and secondary. Tuned coupled circuits of this type are often called "band-pass filters," since fairly uniform amplification can be secured over a wider band of frequencies than is possible when the secondary alone is tuned.[3] The use of several

[3] For a treatment of these circuits based on filter theory the reader is referred to an article by A. J. Christopher, Transformer-coupling Circuits

tuned circuits coupled in cascade will accentuate the band-pass characteristics.

The value of M in the circuit of Fig. 163 is usually made equal to, or slightly greater than, the critical value. In addition to the flatter top, the sides of the curve fall off more rapidly than with a single tuned circuit so that the selectivity is much better.

When the primary and secondary circuits are both tuned alike so that $\omega L_1 = 1/\omega C_1$ and $\omega L_2 = 1/\omega C_2$, (22) becomes

$$A_v = \frac{\mu \dfrac{M}{C_2}}{\sqrt{(R_1R_2 + \omega^2M^2)^2 + \left(\omega L_1 R_2 + r_p \dfrac{R_1R_2 + \omega^2M^2}{\omega L_1}\right)^2}} \tag{23}$$

With fixed coupling between primary and secondary the width of the response curve becomes greater as the frequency increases, which causes a progressive reduction in selectivity in much the same manner as with a single tuned circuit. The gain is also lower than with a single tuned circuit. Tuned coupled circuits are admirably suited for the intermediate-frequency amplifier of a superheterodyne receiver. Here the frequency of the band to be amplified is fixed so that the problem of varying selectivity is not encountered.

99. Cascade Amplifiers.—If two or more identical stages of amplification are connected in cascade the overall gain has been previously shown to be

$$A_T = A_v^n$$

where n is the total number of stages. If the amplification per stage is expressed in decibels, the total gain will be

$$A_T = nA_{db}$$

These expressions presume that the various stages do not react upon each other, which is not always the case in practice because of small unavoidable couplings between input and output circuits. The response curve of a multistage amplifier composed of n identical stages is readily obtained from the curve of a single stage by raising the ordinates of the latter to the nth power.

---

for High-frequency Amplifiers, *Bell System Tech. Jour.*, vol. XI, p. 608, October, 1932.

It is possible to improve both the selectivity and the fidelity of a tuned radio-frequency amplifier by increasing the number of stages used, provided the tuning of each is made broader as their number is increased. This is illustrated in Fig. 164 using the response curves of Fig. 160. Curve A shows the relative amplification of a four-stage amplifier, each stage having the constants of the upper curve of Fig. 160. Curve B is the curve for $M = 30$ in the same figure. The necessity for broader tuning per stage in multistage amplifiers in order to avoid too great an impairment

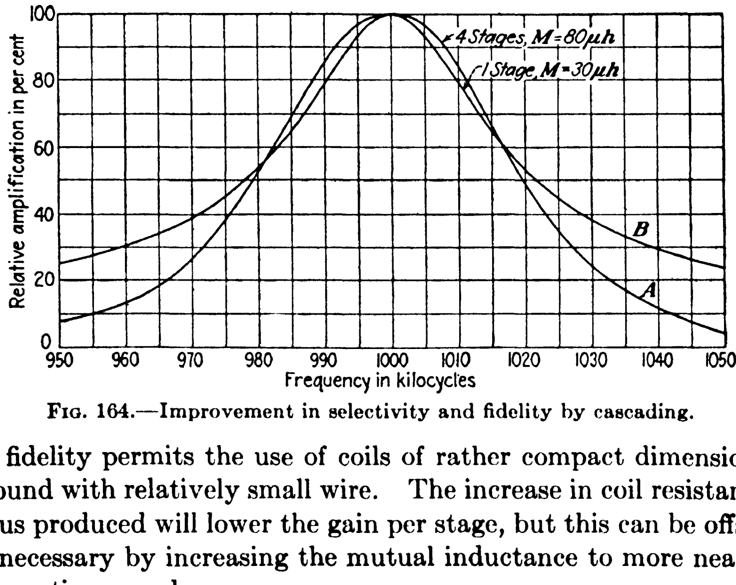


FIG. 164.—Improvement in selectivity and fidelity by cascading.

in fidelity permits the use of coils of rather compact dimensions wound with relatively small wire. The increase in coil resistance thus produced will lower the gain per stage, but this can be offset if necessary by increasing the mutual inductance to more nearly the optimum value.

At frequencies sufficiently remote from resonance where the gain per stage is less than unity, a cascade amplifier acts as an attenuator of the signal. This requires that the amplifier be well shielded in order to prevent stray pick-up by the wiring from reducing this attenuation. Thus, a few inches of exposed wire running from the output stage to the grid of the detector tube may serve as an antenna and have a voltage induced in it from an interfering powerful local station which would be greater in magnitude than the voltage produced by these same signals after passing through the amplifier. The coils, tubes, and tuning condensers of modern receiving sets are all individually shielded so as to prevent oscillations caused by stray couplings between

stages. These necessary precautions also serve to prevent interference from stray pick-up.

100. Methods of Avoiding Oscillation.—All of the preceding types of circuits where triodes are used are almost certain to be unstable at the higher values of frequency because of the feed-back through the coupling capacitance between the grid and plate electrodes. The tendency to oscillate increases as the amplification per stage is increased. With screen-grid tubes the capacitance between control-grid and plate is reduced to a very small value, as will be seen from Fig. 165. The capacitance C_{gp} is broken up in effect into two series condensers with the mid-point grounded to the filament, so far as alternating-current potentials are concerned. Feed-back of amplified output energy through the tube is thereby reduced to the point where stable operation can be obtained at the shortest wave lengths now used. Capacitive coupling between grid and plate leads external to the tube must be carefully avoided by the use of adequate shielding.

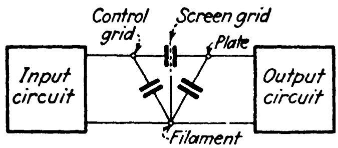


Fig. 165.—Elimination of coupling between input and output circuits by means of a screen-grid tube.

Radio-frequency amplifiers using triodes must be provided with some means of preventing undesired oscillations from occurring. Methods of combating instability are of three general types:

1. By introducing sufficient resistance into the circuit to annul the negative input resistance of the tube.
2. By adjusting the magnitude or phase of the load in the plate circuit so that the input resistance of the tube is not sufficiently negative to cause oscillation.
3. By arranging the circuit so as to form some type of a balanced impedance bridge with the input and output terminals of the tube forming two pairs of diagonally opposite points.

The most common method of the first type is to insert a resistance of about a thousand ohms or less in series with the grid of the tube. In a tuned amplifier designed to cover a range of frequencies this resistance must be large enough to secure stability at the highest frequency. At the lower frequencies where the tendency to oscillate is less, this resistance is much larger than necessary, which results in a loss of amplification and selectivity

in this region. In a number of instances where this method was used in commercial receiving sets, only a portion of the stability was secured in this manner; the remainder was obtained by utilizing some stray coupling between the coils or tuning condensers so that a bridge circuit was produced in effect.

Another method applied an adjustable positive bias to the grid of the tube by connecting the grid-return lead to the arm of a potentiometer connected across the filament-heating battery. This shunted the tuned input circuit with an adjustable positive resistance due to the flow of grid current.

The second type of oscillation control can be secured by keeping the reactance of the load in the plate circuit capacitive so that R_g remains positive. This can be accomplished with the circuits of Fig. 162 by the proper choice of C_m. The plate-circuit impedance may be allowed to become inductive by a small amount as long as R_g does not, as a result, become sufficiently negative to cause instability. This method of oscillation control usually entails an appreciable sacrifice in amplification.

101. Neutralizing Circuits.—With triodes, the third type of oscillation control is the most satisfactory method, as the feed-back voltage is balanced out without impairing the performance of the amplifier. The other two methods merely alleviate the effects of feed-back. In Chap. VIII it was pointed out that regenerative effects caused by feeding back a portion of the amplified output energy to the input circuit through the capacitance between the grid and plate was the equivalent of a negative resistance looking into the grid-filament terminals of the tube. Another point of view which is more helpful in the understanding of neutralizing circuits is that the plate swings through large variations in voltage, and in so doing induces an alternating voltage on the grid through the capacitance between them. To avoid the effects of this, a similar voltage of the opposite phase must be induced on the grid from some source. All attempts to neutralize must attend to both the phase and magnitude of the voltage fed back into the grid. The methods used are various adaptations of alternating-current impedance-bridge circuits having conditions of balance which are independent of the frequency.

One form of bridge circuit due to C. W. Rice is shown in Fig. 166, (*a*) being the actual circuit and (*b*) the electrical equiva-

lent with the tube electrodes omitted. The filament terminal of the tube, instead of being connected to the lower end of the input coil, is connected to an intermediate point which divides the inductance into two parts, L_a and L_b. In receiving sets L_a and L_b are usually made equal. The lower terminal n is connected to the plate through a small balancing condenser C_n. The terminals g and n of the input circuit and f and p of the output circuit constitute two pairs of diagonally opposite points of a bridge, as shown in (b). If the bridge is balanced, no voltage can

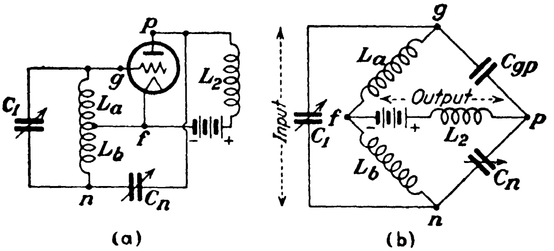


FIG. 166.—Rice neutralizing circuit.

exist across the input terminals gn, owing to a voltage between the output terminals pf. Another way of viewing the circuit is to regard the energy fed back through C_{gp} as being opposed in phase by that which flows through C_n. The conditions for a balance are

$$\frac{L_a}{L_b} = \frac{C_n}{C_{gp}} \tag{24}$$

This balance is independent of the frequency as given by (24) only when the coupling between L_a and L_b is substantially unity, since L_a is shunted by the input capacitance of the tube. With perfect neutralization the input capacitance is constant and independent of the load in the plate circuit. A high-frequency parasitic oscillation may sometimes occur with this type of circuit, which will render it practically inoperative as an amplifier. These oscillations result from subsidiary resonant circuits involving tube capacitances and lead inductances. A small condenser of about the size of C_n shunted across L_2 will often prevent such parasites in receiving circuits. The Rice circuit is commonly used in neutralizing radio-frequency power amplifiers in transmitting sets.

Another form of balancing circuit due to L. A. Hazeltine, known as the neutrodyne circuit, is illustrated in Fig. 167. This circuit applies the same principle to the output circuit as the previous method did to the input. The conditions for a balance are the same as (24). The coupling between L_a and L_b should again be approximately unity if the circuit is to remain balanced for a wide range of frequencies with a fixed adjustment of C_n, as L_a is shunted by the output impedance of the tube. Closer coupling

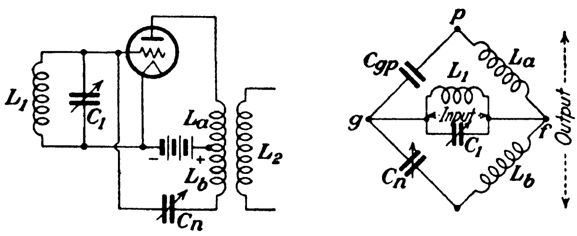


FIG. 167.—Neutrodyne circuit.

between these two coils can be more readily secured than with the input coil in the Rice circuit, since the primary winding of the output transformer is composed of comparatively few turns of fine wire. Another advantage possessed by the neutrodyne circuit in receiving sets is that one set of plates of the tuning condenser C_1 is at filament or ground potential. This enables the rotors of these condensers to be mounted on a common shaft without the need of insulating bushings or couplings when several stages are used in cascade. The neutrodyne circuit was used quite extensively in broadcast receiving sets prior to the general introduction of screen-grid tubes in 1929. An early form of this circuit had the neutralizing condenser C_n connected to a tap at some intermediate point in L_2, thus dispensing with the coil L_b. The secondary coil L_2 must be wound with respect to L_a so that in tracing from the plate connection of L_a to the grid connection of L_2 we continue around the coils in the same direction. Lack of tight coupling between L_a and L_2 with this arrangement makes it more difficult to secure stability over a wide range of frequencies with a fixed adjustment of C_n.

A circuit wherein all four of the bridge arms are condensers is shown in Fig. 168. The grid-plate and grid-filament capacitances

of the tube are used as a pair of ratio arms. The conditions for a balance are

$$\frac{C_n}{C_a} = \frac{C_{gp}}{C_{gf}} \tag{25}$$

The value of C_a is made about 100 $\mu\mu$f, which requires a value of C_n somewhat larger in size than the neutralizing condensers of the

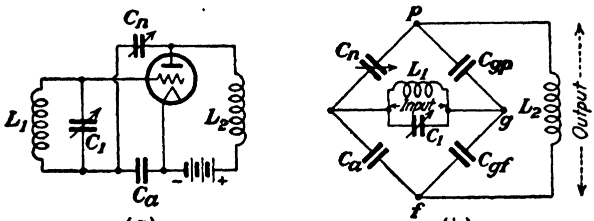


FIG. 168.—Capacitance-bridge neutralizing circuit.

preceding circuits. In order to prevent the accumulation of electrons on the grid which may cause the tube to "block," C_a is usually shunted by a ¼-megohm resistance. The distributed capacitance of a suitable choke coil operated above its resonant frequency can be substituted for C_a.

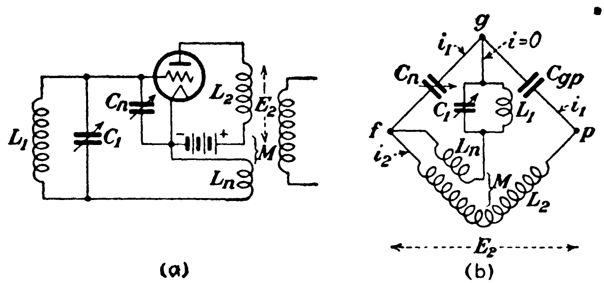


FIG. 169.—Neutralizing circuit employing the principle of a mutual-inductance bridge.

Another form of circuit involving the principle of a mutual-inductance bridge is illustrated in Fig. 169. From the equivalent circuit the following relations are obtained:

$$i_1 = \frac{E_2}{\dfrac{1}{j\omega C_n} + \dfrac{1}{j\omega C_{gp}}} = \frac{-E\omega^2 C_n C_{gp}}{j\omega C_n + j\omega C_{gp}} \tag{26}$$

$$i_2 = \frac{E_2}{j\omega L_2} \tag{27}$$

If the current i in the input circuit due to the output voltage E_2 is to be zero, then

$$\frac{i_1}{j\omega C_n} = ji_2\omega M \tag{28}$$

Substituting the values of i_1 and i_2 from (26) and (27) in (28), we get

$$\frac{M}{L_2} = \frac{C_{gp}}{C_{gp} + C_n} \tag{29}$$

which is the condition for a balance. The neutralizing coil L_n could have been inserted in the grid circuit instead of the grid-return lead, which would then allow one side of the tuning condenser to be at ground potential for convenience in cascading.

Modern receiving sets no longer use triodes in the radio-frequency amplifier but employ screen-grid tubes instead which do not require neutralizing. When triodes are used in the radio-frequency amplifiers of transmitting stations, neutralizing circuits have to be used to secure stability. Screen-grid tubes are also being used extensively in these installations in the low-power stages. It is difficult to design screen-grid tubes for outputs much above 500 watts so that triodes must be used when higher outputs are required.

102. Neutralizing Adjustments.—The neutralizing adjustment in receiving sets is made by tuning the amplifier to strong signal, preferably in the high-frequency tuning range, and then disconnecting the heater or filament connection of the tube to be neutralized. This destroys the repeater action of the tube and converts the stage into its equivalent electrical network. If the signal is of sufficient strength it will pass through the capacitive network of the dead tube with enough energy so as to be heard in the loud-speaker. The neutralizing condenser is then adjusted until the signal disappears. The filament is then lighted and the procedure is repeated with the next stage.

When stray electromagnetic or electrostatic couplings are present, or if the couplings between L_a and L_b in the preceding circuits are appreciably less than unity, the value of balancing capacitance required may vary with the tuning, so that when neutralized at one frequency, the stage may be sufficiently unbalanced at some other frequency to cause oscillations. In this case

a compromise adjustment of C_n must be found, if possible, which will hold the stage out of oscillation for the entire tuning range.

The amplifier stages of transmitting sets are usually provided with a low-reading radio-frequency ammeter in the output tank circuit which serves as the galvanometer in the equivalent bridge circuit. The neutralizing condenser is adjusted for zero deflection of the ammeter when radio-frequency excitation is applied to the grid of the tube to be neutralized. The plate voltage in this case is usually removed, instead of opening the filament circuit to prevent the amplifying action of the tube.

Problems

1. A tuned radio-frequency transformer of the type of Fig. 157 is to be designed to cover the broadcast band of 540 to 1500 kc. The maximum capacitance of the tuning condenser is 350 $\mu\mu$f and its minimum capacitance is 15 $\mu\mu$f. The input capacitance of the tube is 20 $\mu\mu$f and the distributed capacitance of the coil and wiring is 5 $\mu\mu$f. What value of secondary inductance is needed? How many turns of No. 30 enameled wire wound 88 turns per inch are needed for the secondary if the coil form is 1.25 in. in outside diameter? What is the highest frequency to which the transformer can be tuned?

2. A tuned radio-frequency transformer is to be designed for aviation use to cover the frequency band from 230 to 500 kc, using the same tuning condenser and other values of capacitance as in Problem 1. What value of secondary inductance is needed? If a coil form 2 in. in diameter is used, how many turns of No. 30 enameled wire are required? What is the highest frequency to which the transformer can be tuned?

3. *a.* The constants of a tuned radio-frequency amplifier stage of type of Fig. 157 are as follows: $\mu = 10$, $r_p = 10{,}000$ ohms, $M = 20$ μh, $L_2 = 250$ μh, $R_2 = 10$ ohms, $f = 10^6$ cycles, and $\omega L_2 = 1/\omega C_2$. What is the voltage amplification per stage?

b. What would be the voltage amplification per stage if M were adjusted to its optimum value? What is the optimum value of M?

4. In Problem 3 with the constants as in *a*, what is the amplification for an interfering signal 50 kc higher than the resonant frequency of 10^6 cycles?

5. The constants of a tuned radio-frequency amplifier stage of the above type designed for use with a type 58 pentode are as follows: $\mu = 1200$, $r_p = 800{,}000$ ohms, $M = 40$ μh, $L_2 = 300$ μh, $Q_2 = 120$, $f = 10^6$ cycles. What is the resonant voltage amplification per stage?

6. In Problem 5, what is the resonant voltage amplification at 550 and 1500 kc, if $Q_2 = 100$ at these frequencies?

7. In Problem 5, what would be the voltage amplification for an interfering signal 50 kc higher than the resonant frequency of 10^6 cycles?

8. A type 34 pentode has a parallel-resonant circuit in its plate circuit. The tube constants are $\mu = 620$, $r_p = 10^6$ ohms. The constants of the coil

are $L = 280\ \mu h$ and $Q = 80$. What is the voltage amplification at 10^6 cycles, assuming the circuit to be tuned to resonance by means of the condenser across the coil?

9. In Problem 8, what would be the voltage amplification for an interfering signal 50 kc higher than the resonant frequency of 10^6 cycles?

10. A stage of an intermediate-frequency amplifier of a superheterodyne is similar to Fig. 163. The primary and secondary coils are identical and both are tuned to 175 kc. The constants of the coils are $L = 6000\ \mu h$, $Q = 100$, $M = 60\ \mu h$. What is the voltage amplification per stage when using a type 58 pentode of the constants given in Problem 5?

CHAPTER X

OSCILLATORS AND RADIO-FREQUENCY POWER AMPLIFIERS

103. Oscillators.—Any amplifying device is capable of generating oscillations if a sufficient portion of the output energy is fed back into the input in the proper phase so as to reinforce the input energy. Thus, if an ordinary carbon-grain transmitter is connected in series with a battery and a telephone receiver and the two instruments placed so that the sound output of the receiver enters the transmitter, then any initial noise picked up by the transmitter will cause the combination to oscillate. This mechanical oscillator is called a "telephone howler," so named because of the character of the sound produced. The carbon-grain transmitter is an amplifying device, in that the electrical output can be much larger than the input in the form of sound energy. The sound waves cause a variable pressure to be exerted on the carbon granules, causing a variation in the electrical resistance of the transmitter which, in turn, controls the flow of a much larger amount of energy from the local battery.

The initial sound impulse which acted upon the transmitter is reinforced by the resulting output of the receiver so that the amplitude of the oscillations continues to increase until a condition of equilibrium is reached. This stable condition is brought about by the increase in losses as the amplitude of oscillation increases, which causes a corresponding reduction in the efficiencies of the two instruments. With constant efficiency, the ratio of output to input energy in the transmitter and receiver would remain constant and the amplitude of the oscillations would increase indefinitely. The frequency of oscillation will be dictated by the resultant resonant frequency of the two diaphragms.

104. Vacuum-tube Oscillators.—A vacuum-tube oscillator may be regarded as a self-excited amplifier, in that a portion of its output energy is used to excite the grid. One form of circuit in which this may be accomplished is shown in Fig. 170. The

required grid-excitation voltage is induced in the coil L_g which is inductively coupled to the coil L_o in the tank circuit. This voltage will be $\pm j\omega MI_L$, depending upon the sign of M, where I_L is the oscillatory current flowing in L_o. It will be shown presently that the sign of M must be negative in order that the circuit may be self-exciting. The magnitude of the grid voltage is much larger than in the case of the amplifiers previously considered and the grid is driven positive by a considerable amount, causing grid current to flow during a portion of the cycle. The value of E_g can be varied either by changing the number of turns on L_g or, if the latter are fixed, by altering the coupling between L_o and L_g.

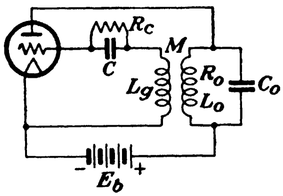


FIG. 170.—Typical oscillator circuit.

Oscillators are operated as Class C devices and are biased negatively to a point considerably beyond cut-off. The necessary grid bias is obtained by means of a resistor or grid leak R_c in series with the grid. The grid current consists of a series of rectified impulses which flow through R_c, so that the biasing voltage will be equal to the average value of this rectified grid current multiplied by the value of the biasing resistance R_c. The grid condenser C shunted across R_c acts as a radio-frequency by-pass to enable the full value of excitation voltage induced in L_g to be impressed across the input terminals of the tube. This condenser does not let the voltage across the resistor vary at a rapid rate, because it supplies charge when the voltage drops a little and absorbs it when the voltage rises. The value of C is not particularly critical and should be large enough so as to have a reactance which is small compared with R_c at the operating frequency. If the capacitance of the grid condenser is made too large, blocking may occur and the oscillations will be periodically started and stopped. The rate at which this condenser discharges is determined by CR_c, the time constant of the circuit. If C is relatively large the condenser will be unable to discharge completely during the negative half cycle in which grid current is zero, with the result that the charge accumulates to the point where oscillations cease because of excessive negative bias. A battery or any other source of constant potential can be used for biasing purposes, but Class C operation requires a bias which exceeds cut-off.

Consequently, when the plate and filament voltages are initially switched on, the plate current will be zero and the circuit will refuse to oscillate until the biasing voltage is reduced to a value less than cut-off. When the circuit is self-biased by means of a resistance in the grid circuit, the initial value of bias is zero, building up automatically to the proper value when oscillations begin.

The initial impulse which starts the oscillations is furnished by switching on the plate voltage, or the filament voltage if the plate voltage is already applied. The growth of plate current i_p through L_o induces a voltage in L_g equal to $M\frac{di_p}{dt}$, and by proper choice of the sign of M a voltage will be applied to the grid which increases i_p. As saturation is approached, di_p/dt becomes less, and finally becomes zero. The increase in the plate current up to this point has been caused by the induced voltage applied to the grid, and when this ceases, the plate current falls. But this decreasing current causes an induced voltage in the opposite direction in the grid circuit, which further decreases the current until cut-off is reached, and the whole cycle starts over again. The frequency of oscillation is dictated by L_o and C_o, which constitute a parallel-resonant circuit connected across the output terminals of the tube.

Maximum power output would presumably be obtained when the impedance of the load was made equal to the internal resistance of the tube. However, the latter varies between extremely wide limits, being infinite during the greater portion of the cycle under Class C operation, so that the load impedance for maximum output bears no simple relation to the plate resistance of the tube, as with Class A operation, where the variation in r_p is small. Best operation will be secured with some particular value of tank-circuit impedance. This impedance is a function of R_o, L_o, and C_o, and if these items are fixed by other considerations the impedance offered by the tank circuit may differ from the optimum value by a considerable amount. As pointed out in connection with audio-frequency amplifiers, it is possible to step up or step down the impedance offered by the load by means of a suitable transformer. Such transformer action can be secured by using the coil in the tank circuit as an autotransformer, as shown in Fig. 171. The ratio of transformation will be approximately

the turns ratio P/S. The frequency of oscillation is governed primarily by the inductance included in S, in conjunction with C_o, while moving the plate tap so as to alter the number of turns included in P changes the impedance of the load as viewed from the tube and has only a minor effect on the frequency. In this way the load impedance offered to the tube may be adjusted to its optimum value.

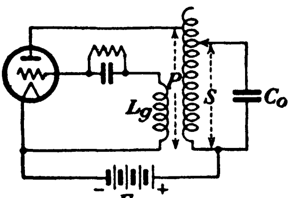


FIG. 171.—Use of tank-circuit inductance as an autotransformer to vary the load impedance.

An approximate treatment of the tuned-plate oscillator that enables some general principles to be deduced which apply to all oscillators may be had from Fig. 172. Applying Kirchhoff's laws to the equivalent circuit and assuming steady-state conditions, we obtain the following equations:

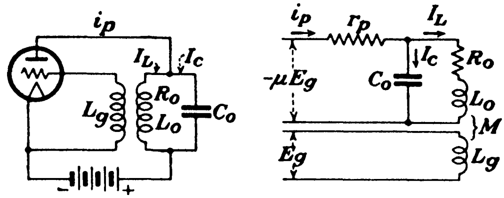


FIG. 172.—Tuned-plate oscillator and its equivalent circuit.

$$i_p = I_L + I_C \tag{1}$$
$$-\mu E_g = i_p r_p + I_L(R_o + j\omega L_o) \tag{2}$$
$$\frac{I_C}{j\omega C_o} = I_L(R_o + j\omega L_o) \tag{3}$$
$$E_g = -j\omega M I_L \tag{4}$$

Substituting (1) and (4) in (2),

$$j\mu\omega M I_L = (I_L + I_C)r_p + I_L(R_o + j\omega L_o) \tag{5}$$

Substituting the value of I_C from (3) in (5),

$$j\mu\omega M I_L = j\omega C_o I_L(R_o + j\omega L_o)r_p + I_L(R_o + r_p + j\omega L_o) \tag{6}$$

from which

$$I_L[(R_o + r_p - \omega^2 L_o C_o r_p) + j\omega(L_o + C_o R_o r_p - \mu M)] = 0 \tag{7}$$

For (7) to be zero the two terms in the brackets must both be zero. Setting the j term equal to zero, we get

$$M = \frac{L_o + C_oR_or_p}{\mu} \tag{8}$$

which gives the minimum value that M can have in order that the circuit may oscillate.

Setting the real term in (7) equal to zero, we get

$$\omega = \sqrt{\frac{R_o + r_p}{L_oC_or_p}} = \omega_o\sqrt{1 + \frac{R_o}{r_p}} \tag{9}$$

where $\omega_o = 1/\sqrt{L_oC_o}$. From (9) it is seen that the frequency of oscillation is affected to some extent by the resistance of the load as well as by r_p of the tube. Consequently, any changes in the

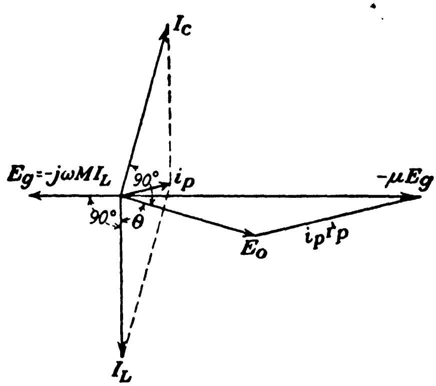


FIG. 173.—Vector diagram of tuned-plate oscillator.

operating conditions of the tube which cause a variation in r_p, such as variations in the plate-supply voltage or filament temperature, will cause slight changes in the frequency of oscillation.

The above treatment neglects the flow of grid current and assumes that the various currents are sinusoidal—conditions that do not exist in practice. Equations (8) and (9) do, however indicate in a general way the relations which exist. The vector diagram of the circuit of Fig. 172 is shown in Fig. 173, again assuming the currents and voltages to be sinusoidal. The grid voltage and the resultant voltage μE_g acting in the plate circuit are 180 degrees out of phase, which accounts for the negative

sign of μE_g in (2). The output voltage across the load is E_o, which is equal to $-\mu E_g$ minus the internal drop $i_p r_p$ within the tube. The current I_L will lag behind E_o by an angle θ, depending upon the ratio of ωL_o to R_o of the coil. The excitation voltage E_g must lag behind I_L by 90 degrees so that the sign of M must be negative in order to produce this condition. The current I_C through C_o will lead E_o by almost 90 degrees, depending upon the losses in this portion of the tank circuit. The plate current i_p will be the vector sum of I_L and I_C. Actually, i_p is badly distorted, being zero for more than half of the cycle, so that the vector representing i_p in Fig. 173 must be regarded as the fundamental component of the plate current.

105. Oscillator Circuits.—The tuned-plate oscillator circuits shown in the preceding diagrams employ series feed, since the tube, tank circuit, and plate-supply voltage are all connected in series. The alternating component of plate current will therefore flow through the source of E_b unless the latter is by-passed by a condenser. This by-pass is absolutely essential when E_b is obtained from a high-voltage direct-current generator because of the high inductance offered by the armature winding. Series feed also causes the tank circuit to be at a high positive potential of E_b volts with respect to the filament which is usually at ground potential.

Shunt feed, as shown in Fig. 174*a*, may be used, which avoids these objections. Additional circuit elements in the form of a radio-frequency choke coil and a blocking condenser C are required. The choke coil excludes radio-frequency currents from the source of plate-supply voltage and at the same time prevents this source from short-circuiting the alternating output voltage E_o. The blocking condenser prevents E_b from being short-circuited by the tank inductance. The value of C is not critical and should be large enough so as to have a value of reactance which is small compared to the tank circuit impedance at the operating frequency.

The Hartley circuit of Fig. 174*b* is very widely used as only a single coil is needed with an intermediate tap brought out to the filament. This divides the tank inductance L_o into two parts, L_p and L_g. The series feed arrangement is rather unsatisfactory, except for small battery operated oscillators, as any capacitance between the filament heating source and E_b will be shunted across

L_p. The filaments of oscillator tubes are often heated by alternating current from a suitable step-down transformer connected to the supply line. The plate voltage E_b is frequently obtained from a rectifier operating from the same source of supply, so that this capacitance, chiefly in the form of capacitance between primary and secondary windings of the filament and rectifier transformers, may be appreciable. Furthermore, the transformer insulation between these windings is subjected to the

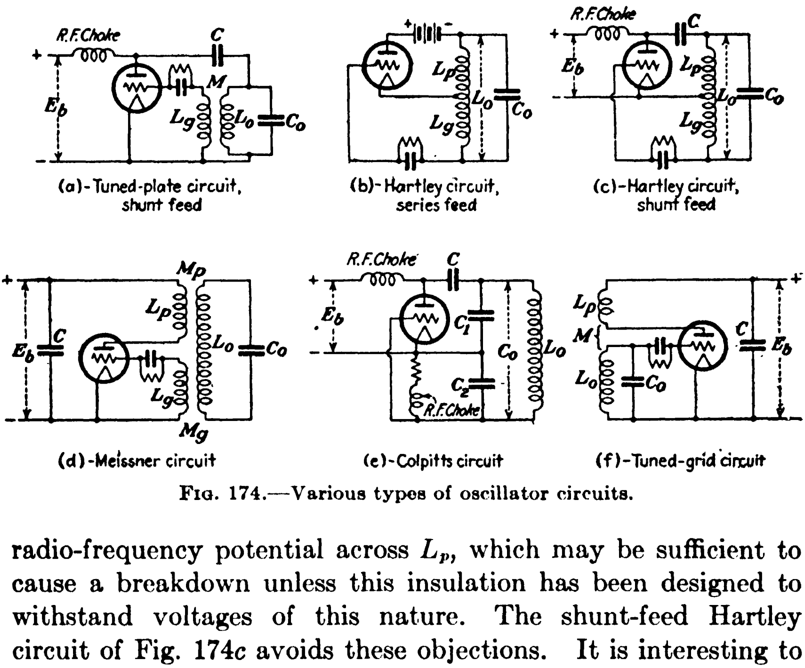


FIG. 174.—Various types of oscillator circuits.

radio-frequency potential across L_p, which may be sufficient to cause a breakdown unless this insulation has been designed to withstand voltages of this nature. The shunt-feed Hartley circuit of Fig. 174c avoids these objections. It is interesting to note the similarity between this circuit and the Rice neutralizing circuit of Fig. 166. As the neutralizing condenser C_n of the latter is increased in size, it begins to function as the blocking condenser C in the shunt-fed Hartley circuit and oscillations begin.

The Hartley circuit becomes the same as the tuned-plate circuit if the lower tap from C_o to the coil is moved up so that C_o is shunted only across L_p; the latter value of inductance now being L_o of the previous diagrams. In practice, all of the tap connections to the coil are usually made adjustable and the one type of circuit evolves into the other as the positions of the various taps are changed.

The Meissner circuit of Fig. 174*d* has the tank circuit inductively coupled to the tube, instead of conductively coupled, as in the previous circuits.

The Colpitts circuit of Fig. 174*e* is similar to the Hartley circuit, except that two condensers C_1 and C_2 in series replace L_p and L_g. The grid-excitation voltage is adjusted by changing the value of C_2, which in this case affects the frequency of oscillation, since the tank capacitance C_o is made up of C_1 and C_2 in series. These two condensers must each have a greater capacitance than the value of C_o used in the previous circuits. However, C_2 also serves as a grid-leak condenser. A radio-frequency choke coil is usually connected in series with the grid leak, since the latter is shunted across a portion of the oscillatory circuit.

The tuned-grid circuit of Fig. 174*f* differs from the tuned-plate circuit only in that the condenser C_o is transferred from L_p to L_g.

In the earlier types of radio transmitters the oscillator was directly coupled to the antenna, the antenna capacitance often serving as C_o. This was very undesirable as any change in the antenna constants due to the wires swinging in the wind, etc., would change the frequency of oscillation. It is now almost a universal practice to use a *master oscillator* operating at low power and to amplify its output by suitable radio-frequency power amplifiers, the output of which is fed to the antenna. The frequency of the master oscillator is usually controlled by the mechanical period of vibration of a quartz crystal whenever the frequency of oscillation must be held to exceedingly close limits, as in modern broadcasting. Crystal oscillators will be discussed later in the chapter.

Two or more tubes may be operated in parallel in any of the above circuits where more power is desired than can be obtained from a single tube. However, it is usually better practice to use a single tube of larger capacity whenever possible as parasitic oscillations are apt to occur where tubes are operated in parallel. These oscillations result from subsidiary resonant circuits involving lead inductances and tube capacitances and are of a much higher order of frequency than the desired oscillations. Such parasitic oscillations may also occur when a single tube is used. The energy represented by these oscillations reduces the output available at the desired frequency, sometimes to only a fraction of its former value, so that they are to be avoided. The insertion

of resistances in locations that will not materially affect the desired oscillations and at the same time will increase the resistance of the parasitic resonant circuits so as to prevent the undesired oscillations from occurring are the usual remedies. A resistance in series with the grid of the tube is usually effective.

Where the combined output of two tubes is required, some form of symmetrical circuit such as the push-pull oscillator of Fig. 175 is generally used. Parasitic oscillations are then more readily avoided than when the tubes are operated in parallel. Push-pull oscillators are very commonly used to obtain extremely high frequencies. They permit the use of very short lead wires and in effect connect the electrode capacitances of the two tubes in series, which enables greater output at higher frequencies to be obtained than could be secured from a single tube. At wave lengths of a few meters the tank circuit consists of merely the tube capacitances in conjunction with the inductances of the lead wires.

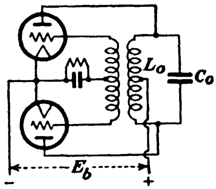


FIG. 175.—Push-pull oscillator circuit.

106. Current and Voltage Relations.—An oscillator converts direct-current energy supplied by the source of E_b in Fig. 176 into alternating-current energy which is furnished to the tank circuit. In order to secure efficiencies of conversion higher than the theoretical maximum of 50 per cent for Class A operation, oscillators and radio-frequency power amplifiers have their grids biased negatively by a sufficient amount so that plate current flows during only a portion of the cycle.

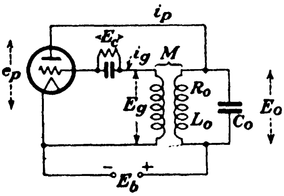


FIG. 176.—Currents and voltages in an oscillator circuit.

The potential e_p of the plate with respect to the filament is at a minimum during the time that the plate current is actually flowing through the tube, so that the power loss within the tube is held to a small value. In this way efficiencies of conversion up to 80 or 90 per cent can be secured. The low value of e_p when i_p is a maximum is due to the impedance drop across the parallel-resonant tank circuit.

Figure 177 shows the instantaneous values and phase relations of the various currents and voltages. The diagram is applicable to either an oscillator or a Class C amplifier, as the latter may be

regarded as a separately excited oscillator, or the former viewed as a self-excited amplifier. Their only differences are in the source of grid excitation. The grid is negatively biased beyond cut-off so that plate current flows for only a portion of the cycle during a time interval equal to $2\theta_1$. The plate current will be badly distorted and will often have a pronounced depression in the top of the wave if the maximum positive swing of E_g is sufficient to reach the region where the dynamic I_p-E_g characteristic bends downward because of excessively large grid current.

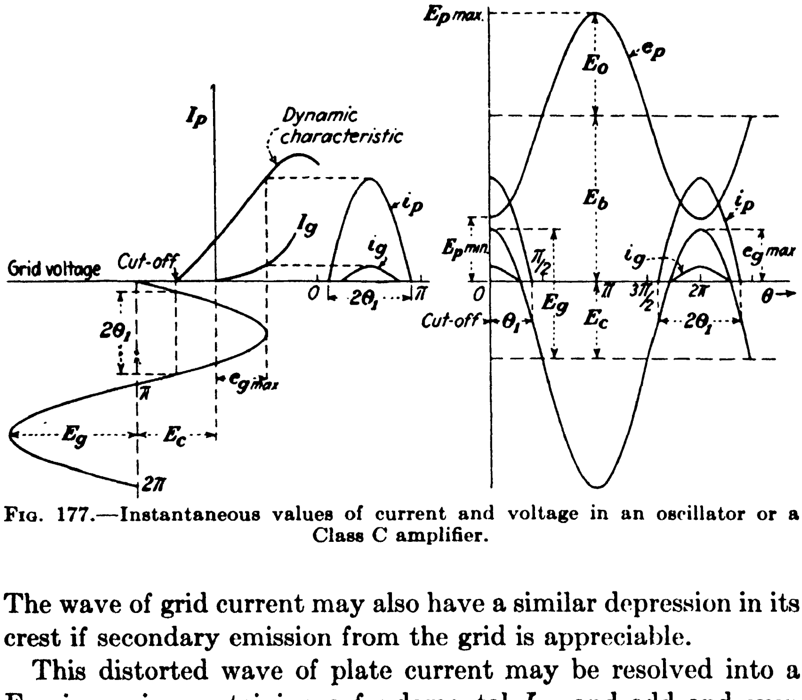


FIG. 177.—Instantaneous values of current and voltage in an oscillator or a Class C amplifier.

The wave of grid current may also have a similar depression in its crest if secondary emission from the grid is appreciable.

This distorted wave of plate current may be resolved into a Fourier series containing a fundamental I_{p_1}, and odd and even harmonics. The tank circuit L_oC_o is resonant to the fundamental, to which it offers a high impedance of the nature of a pure resistance. But the impedance offered to the plate-current harmonics diminishes rapidly with the order of the latter, so that the voltage drop E_o across the tank circuit is very nearly sinusoidal in shape. The instantaneous plate voltage e_p will therefore be the algebraic difference between the plate-supply voltage E_b and the drop E_o across the load.

When the plate current is a maximum, the drop across the load is also a maximum and the plate voltage e_p will have its minimum value at this point. The power loss in the tube is equal to the product of e_p and i_p, averaged over a complete cycle. It is evident from a study of Fig. 177 that this loss can be kept small by limiting the flow of plate current to the brief interval of time that e_p has a low value. The power furnished by the source of plate-supply voltage will be the product of E_b and I_b, the latter being the value of i_p averaged over a complete cycle. The output to the tank circuit will be the input E_bI_b minus the losses within the tube. In the case of an oscillator the grid-excitation losses must be deducted from the energy furnished to the tank circuit to obtain the net output. In a Class B or C amplifier these losses are supplied by the preceding stage. The power furnished to the tank circuit at the fundamental frequency will be the effective value of E_o multiplied by the effective value of the fundamental component of i_p, since the tank-circuit power factor is unity for the fundamental. The impedance of the tank circuit progressively diminishes in magnitude for frequencies above the fundamental and is chiefly composed of condensive reactance, the apparent resistance being small, so that the power factor of the load as well as the voltage across it become smaller for the higher harmonics. Since the power dissipated in an impedance will be $EI \cos \phi$, the power absorbed by the tank circuit will be small for frequencies other than the fundamental.

As already mentioned, E_o will be very nearly a sine wave so that the currents through the coil and condenser will also be of good wave-form. These two currents will be approximately equal in magnitude and nearly 180 degrees out of phase, as shown in the vector diagram of Fig. 173, so that they both may be large in comparison with i_p. As the effective value of Q of the tank coil is increased the circulating current in the tank circuit becomes still larger compared to i_p. This circulating current represents a dissipation of energy in the coil—assuming negligible losses in the condenser—and since the energy to maintain the oscillations is supplied at a different instantaneous rate in the form of series of impulses, some flywheel effect is necessary. Experience has shown that circuits having less than twice as much energy stored in them as they dissipate per cycle tend to be unstable.

When the current through C_o is zero, the voltage across it is a maximum and equal to E_o. At this instant all of the energy is stored in the condenser and is equal to $\frac{1}{2}C_oE_o^2$ joules. The maximum value of the current through C_o is $2\pi fC_oE_o$. If the power lost in the circuit is W watts, the joules lost per cycle are W/f. Dividing the energy stored by the energy lost per cycle gives

$$\frac{\text{Energy stored}}{\text{Energy lost per cycle}} = \frac{\frac{1}{2}C_oE_o^2}{W/f} = \frac{fC_oE_o^2}{2W} = \frac{2\pi fC_oE_o^2}{4\pi W} = \frac{E_oI_o}{4\pi W}$$
$$= \frac{EI}{2\pi W} \tag{10}$$

where E and I are effective values. Therefore, the ratio of the energy stored to the energy dissipated is $\frac{1}{2\pi}$ times the ratio of the volt-amperes in the tank circuit to the watts. To avoid unstable operation in oscillators, the ratio given by (10) should not be less than 2, so that the circulating volt-amperes should be at least 4π times the total power output in watts, or

$$\frac{EI}{W} > 4\pi \tag{11}$$

Since EI/W is approximately equal to $\omega L_o/R_o$, which is the effective value of Q for the coil, it follows that Q should not be less than about 12 if erratic operation is to be avoided. Values of Q greater than this, while satisfactory, result in larger circulating currents in the tank circuit, requiring coils and condensers of greater current-carrying capacity and hence greater cost. Most of the effective resistance R_o in the tank coil is the coupled resistance introduced by the useful load, so that Q is determined primarily by the reflected resistance, assuming the load to be inductively coupled to the tank coil. The resistance of the coil itself should be as small as possible.

It will be noted that the grid-excitation voltage E_g is at its positive maximum value when e_p is a minimum. The minimum plate voltage should not be allowed to fall below the value of $e_{g_{max}}$, if excessive grid current is to be avoided. The plate is located farther from the filament than the grid, so that if these two electrodes both have the same positive potential with respect to the filament the grid current may become comparable to the

plate current, resulting in large losses in the grid circuit. For this reason, the value of $e_{g_{max}}$ is usually limited to about 80 per cent of $E_{p_{min}}$.

107. Circuit Calculations.[1]—In designing the circuits for an oscillator or a power amplifier, the given data will include the frequency, the type of tube to be used, and the plate-supply voltage. The minimum plate voltage $E_{p_{min}}$ and maximum positive value of the grid voltage $e_{g_{max}}$ are then selected, with the restriction that the latter must not exceed the former in value and is usually in the neighborhood of 80 per cent of $E_{p_{min}}$. The portion of the cycle $2\theta_1$ during which plate current is allowed to flow is next selected. This will vary from 180 degrees in the case of a Class B amplifier to perhaps as low as 60 degrees for Class C operation. Oscillators are always operated Class C. In order to determine the best possible conditions, it is necessary to assume several values of one of the independent variables while the others are kept constant and then repeat for successive values of the others.

If θ is measured from the point of maximum grid voltage in Fig. 177, the expressions for the instantaneous grid and plate voltages will be

$$e_g = -E_c + E_g \cos \theta \tag{12}$$

$$e_p = E_b - E_o \cos \theta \tag{13}$$

Cut-off may be assumed to occur when

$$-\mu e_g = e_p \tag{14}$$

and if this occurs at the angle θ_1, substituting (12) and (13) in (14) gives us

$$\mu E_c - \mu E_g \cos \theta_1 = E_b - E_o \cos \theta_1$$

and the value of C bias required will be

$$E_c = \frac{E_b}{\mu} + \left(E_g - \frac{E_o}{\mu}\right) \cos \theta_1 \tag{15}$$

From Fig. 177, it will be seen that this value of bias voltage is also equal to

[1] A comprehensive treatment of this subject is given in a series of articles by D. C. Prince, Vacuum Tubes as Power Oscillators, *Proc. I.R.E.*, vol. 11, pp. 275, 405, 527, June, August, October, 1923.

$$E_c = E_g - e_{g_{max}} \tag{16}$$

where $e_{g_{max}}$ is the maximum positive value of the grid voltage. Equating (15) and (16)

$$E_g - e_{g_{max}} = \frac{E_b}{\mu} + \left(E_g - \frac{E_o}{\mu}\right)\cos\theta_1$$

$$E_g(1 - \cos\theta_1) = \frac{E_b}{\mu} - \frac{E_o}{\mu}\cos\theta_1 + e_{g_{max}}$$

Substituting for E_o in the above equation

$$E_o = E_b - E_{p_{min}} \tag{17}$$

we get

$$E_g(1 - \cos\theta_1) = \frac{E_b}{\mu} - \frac{E_b - E_{p_{min}}}{\mu}\cos\theta_1 + e_{g_{max}}$$

or

$$E_g = \frac{E_b}{\mu} + \frac{1}{1 - \cos\theta_1}\left(\frac{E_{p_{min}}\cos\theta_1}{\mu} + e_{g_{max}}\right) \tag{18}$$

which is the maximum amplitude of the grid-excitation voltage required. Knowing the value of E_g from (18), the biasing voltage E_c can be obtained from (16).

Corresponding pairs of plate and grid voltages can then be computed for increments of 5 or 10 degrees over the time interval $2\theta_1$ during which plate current flows. Since the various current and voltage waves are symmetrical on either side of the y axis, it is only necessary to do this from zero to θ_1. A suitable table for this purpose is as follows:

TABLE I

Given data:	Assumed values:	Computed values:
Tube..................	$E_{p_{min}}$..................	E_g............ eq. (18)
μ....................	$e_{g_{max}}$..................	E_o............ eq. (17)
E_b....................	θ_1....................	E_c............ eq. 16)

1. θ	0°	10°	20°	30°	40°	θ_1
2. $\cos\theta$	1	0.9848	0.9397	0.8660	0.7660	......
3. $E_o\cos\theta$	......	......	......	......	......	......
4. $e_p = E_b - E_o\cos\theta$	......	......	......	......	......	......
5. $E_g\cos\theta$	......	......	......	......	......	......
6. $e_g = E_g\cos\theta - E_c$	......	......	......	......	......	......
7. i_p	y_0	y_1	y_2	y_3	y_4	0
8. i_g	......	......	......	......	......	0
9. $i_p\cos\theta$	y'_0	y'_1	y'_2	y'_3	y'_4	0
10. $i_g\cos\theta$	......	......	......	......	......	0

The values of plate and grid currents in lines 7 and 8 are obtained from the static characteristics of the tube for the com-

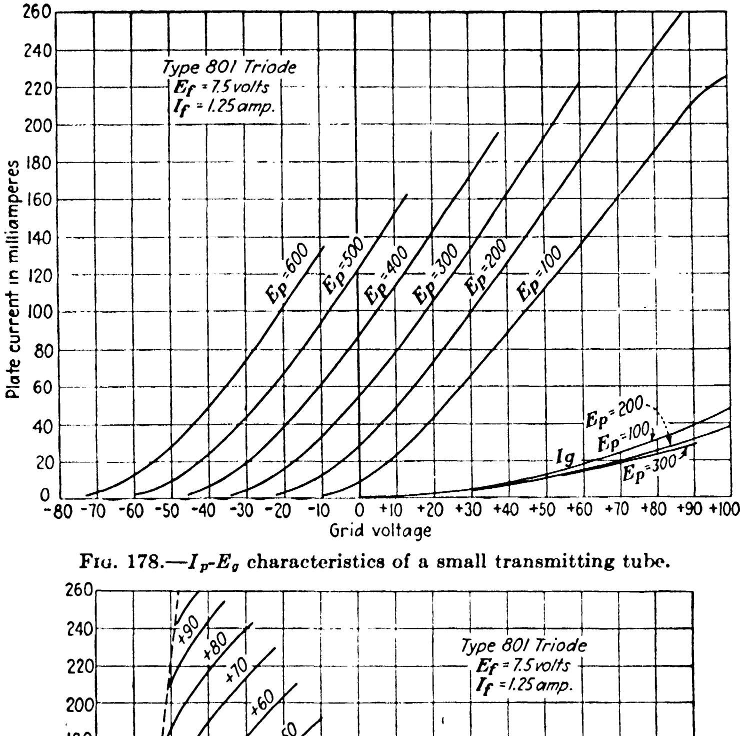


FIG. 178.—I_p-E_g characteristics of a small transmitting tube.

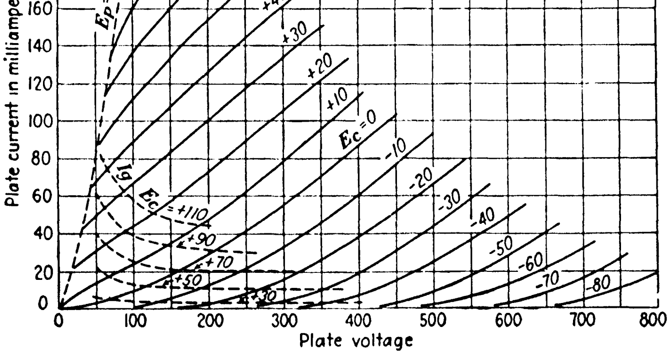


FIG. 179.—I_p-E_p characteristics of a small transmitting tube.

puted instantaneous values of e_p and e_g in lines 4 and 6. Either the I_p-E_g or the I_p-E_p characteristics may be used. These

two families of curves are shown in Figs. 178 and 179 for a typical small transmitting tube. The portion of the characteristic curve for positive values of grid voltage are of particular importance in the case of oscillators and power amplifiers. Most of the items in the preceding table will be for positive values of grid voltage. The grid-current characteristics are also needed in order to determine the power required by the grid circuit.

The direct-current component of plate current I_b will be the average value of i_p over a complete cycle, which will be equal to the area under the curve divided by the base, zero to 2π. Since the wave of current is symmetrical on either side of the y axis the average ordinate may be found with less labor by dividing the area of half the wave by π. The area under the curve of Fig. 180 may be found by means of either Simpson's rule or the trapezoidal

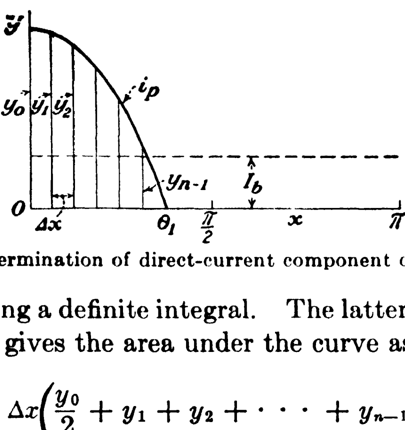


FIG. 180.—Determination of direct-current component of plate current.

rule for evaluating a definite integral. The latter rule is the more convenient and gives the area under the curve as

$$\text{area} = \Delta x\left(\frac{y_0}{2} + y_1 + y_2 + \cdots + y_{n-1} + \frac{y_n}{2}\right) \tag{19}$$

where Δx is the interval between ordinates and is taken as 10 degrees or $\pi/18$ in Table I. The values of the ordinates are obtained from line 7, and since $y_n = 0$, the value of I_b will be

$$I_b = \frac{\text{area}}{\pi} = \frac{1}{\pi} \times \frac{\pi}{18}\left(\frac{y_0}{2} + y_1 + y_2 + \cdots + y_{n-1}\right)$$
$$= \frac{1}{18}\left(\frac{y_0}{2} + y_1 + y_2 + \cdots + y_{n-1}\right) \tag{20}$$

If the interval in line 1 had been 5 degrees, the coefficient of (20) would have been $\frac{1}{36}$, or for any other interval of Δx degrees, it would have been $\Delta x/180$.

The direct-current component of grid current I_c can be found in a similar manner by substituting as ordinates the items of line 8 in (20).

The plate current of Fig. 177 can be resolved into a Fourier series of the form

$$i_p = I_b + I_{p_1} \cos \theta + I_{p_2} \cos 2\theta + \cdots + I_{p_n} \cos n\theta \qquad (21)$$

where I_b is direct-current component just determined, and I_{p_1}, I_{p_2}, I_{p_n} are the maximum values of the fundamental, second harmonic, and nth harmonic, respectively.

To obtain the value of I_{p_1}, multiply both sides of (21) by $\cos \theta d\theta$ and integrate the resulting equation between the limits zero and 2π.

$$\int_0^{2\pi} i_p \cos \theta d\theta = I_b \int_0^{2\pi} \cos \theta d\theta + I_{p_1} \int_0^{2\pi} \cos^2 \theta d\theta + I_{p_2} \int_0^{2\pi} \cos 2\theta \cos \theta d\theta + \cdots + I_{p_n} \int_0^{2\pi} \cos n\theta \cos \theta d\theta$$

All of the terms on the right-hand side of the above equation, except the second, reduce to zero, leaving

$$\int_0^{2\pi} i_p \cos \theta d\theta = \pi I_{p_1}$$

or

$$I_{p_1} = \frac{1}{\pi} \int_0^{2\pi} i_p \cos \theta d\theta = \frac{2}{\pi} \int_0^{\pi} i_p \cos \theta d\theta \qquad (22)$$

The definite integral of the right-hand side of (22) is evidently the area under the curve $y = i_p \cos \theta$ between the limits zero and π, so that if the items in line 9 of Table I are plotted, a curve would be obtained similar in appearance to Fig. 180. The area under this curve, from the trapezoidal rule of (19), with Δx equal to 10 degrees, is

$$\int_0^{\pi} i_p \cos \theta d\theta = \frac{\pi}{18}\left(\frac{y'_0}{2} + y'_1 + y'_2 + \cdots + y'_{n-1}\right)$$

and the maximum amplitude of the fundamental will be given by

$$\begin{aligned} I_{p_1} &= \frac{2}{\pi} \times \frac{\pi}{18}\left(\frac{y'_0}{2} + y'_1 + y'_2 + \cdots + y'_{n-1}\right) \\ &= \frac{1}{9}\left(\frac{y'_0}{2} + y'_1 + y'_2 + \cdots + y_{n-1}\right) \qquad (23) \end{aligned}$$

If 5-degree intervals had been used in Table I, the coefficient of (23) would have been $\frac{1}{18}$.

The maximum amplitude of the fundamental component of the grid current I_{g_1} can be obtained in a similar manner by substituting the items of line 10 in (23). The preceding relationships have been based on the series-fed circuit of Fig. 171, but it is evident that they also apply to the case of shunt feed.

108. Power Relations.—The direct-current power supplied to the circuit from the source of E_b will be

$$P_{\text{input}} = E_b I_b \tag{24}$$

The power output to the tank circuit at the fundamental frequency will be

$$P_{\text{tank}} = \frac{E_o}{\sqrt{2}} \times \frac{I_{p1}}{\sqrt{2}} = \frac{E_o I_{p1}}{2} \tag{25}$$

as the impedance of the tank circuit is of the nature of a pure resistance R_b at resonance. The required value of R_b is evidently

$$R_b = \frac{E_o}{I_{p1}} \tag{26}$$

and will be related to the other circuit constants by

$$R_b = \frac{R_o^2 + \omega^2 L_o^2}{R_o} = \frac{L_o}{C_o R_o} \tag{27}$$

where R_o is the apparent resistance of the tank coil and includes coupled resistance introduced by the useful load which is either inductively or conductively coupled to the tank coil. In the case of an oscillator R_o will also include the reflected resistance from the grid circuit due to the grid-excitation losses.

The useful power output of an oscillator will be the power furnished to the tank circuit minus the input to the grid. In a power amplifier the grid-excitation losses are furnished by the source of separate excitation. The losses in C_o and the I^2R loss in the tank coil itself must also be deducted if they are appreciable.

The losses in the grid circuit will be composed of the power lost in the grid leak R_c and the power lost within the tube owing to the flow of grid current. This power is supplied at the fundamental frequency from the tank circuit in the case of an oscillator, or from the source of excitation in the case of a power amplifier, and will be

$$P_{\text{grid input}} = \frac{E_g I_{g1}}{2} \tag{28}$$

The power lost in the grid leak is

$$P_c = I_c^2 R_c = E_c I_c \tag{29}$$

so that the power consumed within the tube, owing to the flow grid current, will be

$$P_g = \frac{E_g I_{g1}}{2} - I_c^2 R_c \tag{30}$$

Since the grid is enclosed by the plate, the heating of the grid by P_g must be radiated by the plate in addition to its own losses, and is therefore added to the plate loss.

Oscillators are practically always self-biased by means of a grid leak and condenser. This enables an oscillator to be self-starting and has the further advantage of automatically increasing the biasing voltage if the grid-excitation voltage should increase. The required value of grid-leak resistance R_c can be obtained from (29) and is given by

$$R_c = \frac{E_c}{I_c} \tag{31}$$

The phenomena of secondary emission from the grid may cause the grid-current characteristics to vary appreciably among tubes of the same type. It depends, among other things, upon the surface conditions of the grid. A very small film of the "getter" material deposited on the grid during the evacuating process may greatly alter its emission characteristics. The effects of secondary emission on the grid current are shown in Fig. 181 and have already been discussed in Sec. 79. If a fairly high positive potential is applied to the grid when the plate potential is also high, the electrons splashed out of the grid may exceed those attracted to it from the filament, causing the grid current actually to reverse its direction of flow and become negative. This is termed "blocking" in an oscillator and usually results in the destruction of the tube. The reversed direction of grid current through the grid leak converts E_c into a positive bias so that the plate current becomes excessively large and the resultant overheating of the plate with the very probable evolution of absorbed gas will cause the tube to be ruined. Blocking is most apt to

occur when high values of grid-leak resistance are used in an attempt to secure high plate efficiencies.

Radio-frequency power amplifiers operating Class B and Class C are usually biased by means of a battery or a generator instead of using a grid leak and condenser. If the latter were used, a failure of the grid-excitation voltage would remove the negative bias and allow the plate current to rise to a value much greater than with normal bias, with probable damage to the tube. Furthermore, in the case where a modulated signal is to be amplified, the biasing voltage E_c would fluctuate with the frequency of modulation, which would be undesirable. When it

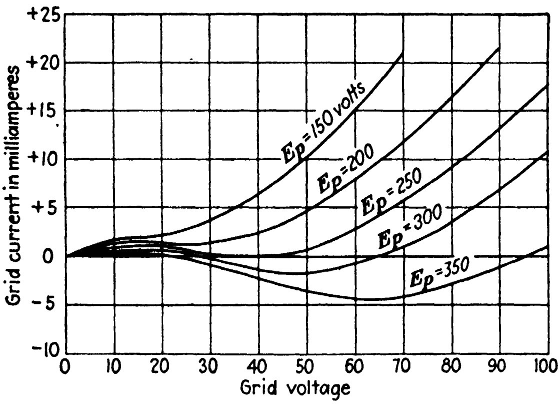


FIG. 181.—Grid-current characteristics caused by secondary emission.

is particularly important to avoid such fluctuations, the source of biasing potential must be selected to have a low internal resistance. For this reason a source such as a thermionic rectifier and filter would be unsatisfactory. Grid leaks are sometimes used to bias a modulated amplifier, where certain advantages result from their use which will be discussed later.

Where a battery is used for biasing purposes the flow of grid current is in a direction such as to charge the battery. The same amount of power will be dissipated in this case as given by (29).

The required value of grid-excitation voltage E_g to produce the assumed operating conditions in the case of an oscillator will be

$$|E_g| = I_L \omega M \tag{32}$$

where I_L is the maximum value of the current flowing in the tank coil of Fig. 176 and is given by

$$|I_L| = \frac{E_o}{\sqrt{R_o^2 + \omega^2 L_o^2}} \tag{33}$$

Where Q of the coil is high, the current in the coil and condenser are approximately the same and will be given with sufficient accuracy for most purposes by

$$|I_L| = |I_C| = E_o \omega C_o = \frac{E_o}{\omega L_o} \tag{34}$$

The plate efficiency of a tube is defined as the ratio of the power output to the tank circuit to the power supplied to the plate, so that

$$\text{Plate efficiency} = \frac{E_o I_{p_1}}{2 E_b I_b} \tag{35}$$

Power tubes are rated on the maximum allowable power that can be dissipated at the plate without excessive heating. This power loss, while treated as though it were due to an internal $i_p^2 r_p$ loss within the tube, actually occurs only at the surface of the plate, and is due to the electronic bombardment of the latter; each electron losing an amount of kinetic energy equal to $\frac{1}{2}mv^2$. If the plate loss is fixed, a moderate improvement in the plate efficiency will greatly increase the useful power output. For example, suppose a tube has a maximum allowable plate dissipation of 50 watts and is operating with a plate efficiency of 50 per cent. The input is therefore 100 watts and the output is 50 watts. But if the plate efficiency could be increased to 80 per cent, the permissible input could be increased to 250 watts, resulting in an output of 200 watts or an increase of 400 per cent, with the same plate loss as before. Consequently, the higher the plate efficiency, the greater the power output from a given tube. As Class C operation results in the highest plate efficiency, this mode of operation is always used when large outputs are of primary consideration.

The power loss within the tube which has to be dissipated at the plate in the form of heat, exclusive of the power loss in the filament, is

$$\text{Tube loss} = E_b I_b - \frac{E_o I_{p_1}}{2} + \frac{E_g I_{g_1}}{2} - E_c I_c \tag{36}$$

This expression may be used to check the assumed operating

conditions from the standpoint of allowable plate dissipation. In the smaller tubes all the heat developed at the plate must be radiated so that the plates are usually carbonized in order to take advantage of the high radiating efficiency of a black body. Plate structures composed of carbon or graphite are also used. About 1000 watts[2] is the maximum amount of power that can be conveniently radiated and for plate losses in excess of this water cooling of the plate must be resorted to.

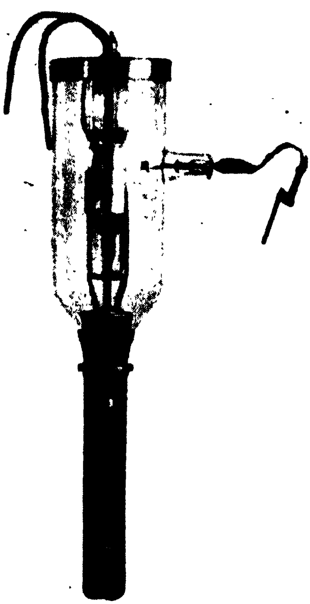


FIG. 182.—Water-cooled tube with portion of plate cut away, showing grid structure.

The plate in water-cooled tubes consists of a seamless copper tube welded into a glass base so as to form a portion of the containing envelope. A typical tube of this type with a portion of the plate cut away to show the grid structure is shown in Fig. 182. The tube shown, type 207, has a maximum output of 20 kw. and a plate dissipation of 10 kw. The filament requires 52 amp. at 22 volts and a plate potential up to 15,000 volts can be used. The plate fits into a water jacket through which the cooling water is circulated in actual contact with the copper plate. As the plate is at a high positive potential with respect to the filament—the latter usually being at ground potential—the cooling water must be supplied by means of rubber hose connections of a sufficient length so that the water columns will have a resistance high enough to prevent undue leakage of current. The maximum plate voltage is almost twice the average value of E_b. The inlet and outlet hose connections are usually coiled together on a cylindrical form located beneath the tube and provide two leakage paths in parallel of about 20 to 30 ft. in length, depending on the

[2] H. E. MENDENHALL, Radiation-Cooled Power Tubes for Radio Transmitters, *Bell Lab. Rec.*, vol. 11, p. 30, October, 1932.

value of the plate potential. The water used must be fairly pure so as to have a high specific resistance. Distilled water circulated in a closed cooling system is often used. Provision must be made to remove the plate voltage from the tube in the event of a failure of the cooling system. The large amount of power dissipated at the plate would quickly produce an excessive temperature rise, which would liberate sufficient gas to destroy the high vacuum within the tube, or even melt the copper plate.

109. Power-amplifier Computations Based upon Approximate Tube Characteristic.—The foregoing analysis, based upon the static characteristics of the tube, involves the choice of direct and alternating grid and plate voltages, from which the load impedance is obtained; whereas in actual operation the load impedance is one of the independent variables and the alternating plate voltage across it is a dependent variable. In order to draw conclusions as to the effects of the various parameters upon the efficiency and output, it is necessary to compute the performance for a sufficient number of conditions so that curves may be drawn from which the best operating conditions may be determined.

An approximate solution due to W. L. Everitt,[3] which reduces the labor involved, is to assume a linear relation between the equivalent grid voltage and plate current. This relation is given by

i_L e_c e_L

$$i_p = g_m\left(e_g + \frac{e_p}{\mu}\right) \tag{37}$$

and neglects the curvature in the characteristic in the vicinity of cut-off, since the mutual conductance g_m is assumed to be constant. As the value of grid-excitation voltage used in Class B and C amplifiers is relatively large, the neglect of the curvature in this region will have only a small effect on the average and fundamental values of the plate current. A more serious objection to the assumption of a linear relationship with larger power tubes is that the effects of saturation at high values of plate current are neglected. This is of more consequence with Class C amplifiers than with Class B, as the operating conditions of the

[3] Optimum Operating Conditions for Class C Amplifiers, *Proc. I.R.E.*, vol. 22, p. 152, February, 1934.

latter are chosen so as to obtain as linear a characteristic as possible.

The voltage and current relations under these assumptions are shown in Fig. 183. Equation (37) is valid only for positive values of $\left(e_g + \frac{e_p}{\mu}\right)$. For negative values of this quantity the current is zero. The impedance of the tank circuit L_oC_o will be assumed to be a pure resistance R_b for the fundamental frequency and to

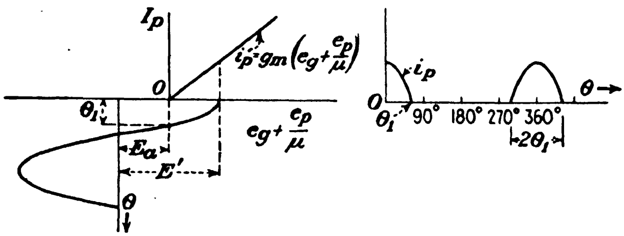


FIG. 183.—Voltage and current relations of a vacuum tube assuming a linear characteristic.

offer negligible impedance to the higher harmonics. The equivalent grid voltage will have an alternating- and a direct-current component defined by

$$E' = E_g - \frac{I_{p_1}R_b}{\mu} \tag{38}$$

$$E_a = E_c - \frac{E_b}{\mu} \tag{39}$$

where E_g = maximum value of the alternating grid-excitation voltage.

E_c = magnitude of the negative bias voltage.

I_{p_1} = maximum value of the fundamental component of the alternating plate current.

From Fig. 183,

$$\cos \theta_1 = \frac{E_a}{E'} \tag{40}$$

$$i_p = g_m(E' \cos \theta - E_a) \tag{41}$$

The plate current can be expressed by (21), as before, the fundamental component of which, from (22) is

$$I_{p_1} = \frac{2}{\pi}\int_0^{\pi} i_p \cos \theta d\theta = \frac{2}{\pi}\int_0^{\theta_1} i_p \cos \theta d\theta \tag{42}$$

since the area under the plate-current curve is zero in the interval from θ_1 to π. Substituting the value of i_p from (41) in (42)

$$
\begin{aligned}
I_{p_1} &= \frac{2g_m}{\pi}\int_0^{\theta_1}(E' \cos \theta - E_a) \cos \theta d\theta \\
&= \frac{E'g_m}{\pi}\left(\theta_1 - \frac{1}{2}\sin 2\theta_1\right) \qquad (43)
\end{aligned}
$$

The direct-current component I_b of plate current will be the average value of i_p over a complete cycle which is

$$
\begin{aligned}
I_b &= \frac{1}{2\pi}\int_0^{2\pi} i_p d\theta = \frac{g_m}{\pi}\int_0^{\theta_1}(E' \cos \theta - E_a)d\theta \\
&= \frac{E'g_m}{\pi}(\sin \theta_1 - \theta_1 \cos \theta_1) \qquad (44)
\end{aligned}
$$

Equations (43) and (44) give the values of the fundamental and the direct-current component of the plate current, respectively, for an amplifier operating either Class B or Class C, under the assumption that the conditions of operation result in a dynamic characteristic which is essentially linear. Where this assumption would result in too great an error the method of Sec. 108 may be used.

The foregoing theory will also apply to screen-grid tubes.[4] With these tubes the plate current will be determined almost entirely by the screen-grid and control-grid voltages, the plate voltage having very little effect. In this case the cut-off voltage will be given approximately by

$$
E_{co} = -\frac{E_{sg}}{\mu_s} \qquad (45)
$$

where E_{sg} is the screen-grid voltage and μ_s is the amplification factor of an equivalent triode with the plate in place of the screen grid.

110. Amplifier-tuning Adjustments.—If the value of E' from (38) is substituted in (44), we get

$$
I_b = \frac{g_m}{\pi}\left(E_g - \frac{I_{p_1}R_b}{\mu}\right)(\sin \theta_1 - \theta_1 \cos \theta_1)
$$

It is evident from this expression, assuming E_g to be constant in amplitude, that I_b will be a minimum when the voltage drop

[4] C. E. FAY, The Operation of Vacuum Tubes as Class B and Class C Amplifiers, *Proc. I.R.E.*, vol. 20, p. 564, March, 1932.

$I_{p_1}R_b$ across the tank circuit is a maximum. This will occur when the impedance of the tank circuit R_b—or, in general, Z_b—is made a maximum. Therefore, a direct-current ammeter in the plate circuit of a radio-frequency power amplifier can be used to indicate when resonance is obtained in the tank circuit. Assuming C_o to be the tuning element, it should be adjusted so that I_b is a minimum.

However, a variable condenser is seldom used in the tank circuit except in low-power amplifiers operating at fairly high frequencies. Where larger amounts of power are handled, C_o is a fixed mica condenser and the tuning adjustments are made by varying the tank inductance. Coarse adjustments are made by moving the leads which terminate in suitable clips so as to include the approximate number of turns needed. A vernier adjustment is usually provided in the form of a heavy copper or aluminum disk which acts as a short-circuited secondary of a single turn, and which can be rotated within the coil L_o. By varying the mutual inductance between this short-circuited turn and L_o a fine adjustment of the tank inductance is obtained for tuning purposes. The position of the clips should be chosen so that the plane of the disk is displaced nearly 90 degrees at resonance from that of the coil, so as to minimize the I^2R loss in the disk.

In this case, since L_o is the variable, maximum impedance will not occur when the power factor of the tank circuit is adjusted to unity, as discussed in Chap. II. This will be seen from the vector diagram of Fig. 34. Consequently, *when the tank circuit is tuned to resonance by varying L_o, the value of the latter should be made a trifle smaller than the value which produces the minimum I_b.* The tank circuit should be adjusted to unity power factor (resonance) rather than maximum impedance, otherwise the maximum plate current will not coincide with the point of minimum plate voltage, and plate losses will be increased. If the effective value of Q of the tank circuit is high, the adjustments for maximum impedance and unity power factor do not materially differ and the circuit can be adjusted for minimum reading of the direct-current ammeter in the plate circuit of the tube, regardless of which tuning element is varied. But when the effective value of Q is comparatively low, as is often the case in practice, minimum I_b can be used as a criterion of resonance only when C_o is the tuning element.

111. Class B Amplifiers.—Class B amplifiers have been defined as those which operate with a negative grid bias such that the plate current is practically zero when the grid-excitation voltage is removed, and in which the power output is proportional to the square of the grid-excitation voltage. They are used in radio-telephone transmitters to amplify the radio-frequency voltage after it has been modulated. The power amplifiers following the modulated stage will therefore all be operated Class B.

With audio-frequency amplification it is necessary to use two tubes in push-pull to avoid distortion, when operating Class B.

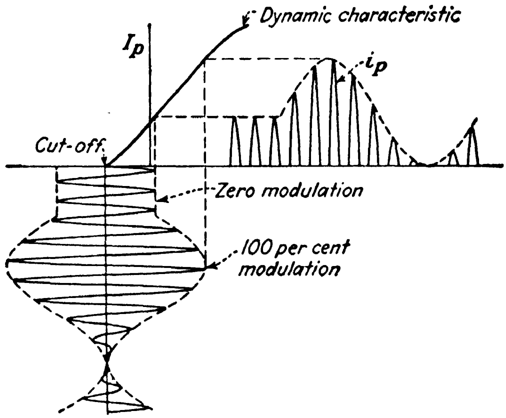


FIG. 184.—Amplification of a modulated radio-frequency wave by a single tube operating Class B.

While the push-pull connection is often used with radio-frequency power amplifiers, a single tube operating Class B will not distort the modulated envelope of voltage applied to the grid, as shown in Fig. 184. Harmonics of the carrier frequency will be present in the plate-current waves, but the audio-frequency envelope will be undistorted if the dynamic characteristic is linear over the operating range. The output tank circuit receives one impulse of varying magnitude per cycle, which causes the current I_o in the tank to rise and fall in accordance with the envelope of the excitation voltage applied to the grid.

When two tubes are used in push-pull, the tank circuit receives two impulses per cycle, and even harmonics of the carrier frequency will presumably be absent in the plate circuit. This greatly reduces the burden imposed upon subsequent tuned cir-

cuits of filtering out these undesired harmonics of the carrier frequency. These harmonics, if allowed to reach the antenna, would be radiated and would cause interference with other stations operating at frequencies of 2, 3, 4, etc., times the carrier frequency of the station in question.

The instantaneous peak output of the tube when the excitation voltage is modulated 100 per cent will be four times the unmodulated output. The continuous power output with this degree of modulation is 1.5 times the output at zero modulation. The unmodulated grid voltage must be low enough so that saturation effects are not appreciable at modulation peaks. A plate efficiency of about 70 per cent may be secured with 100 per cent modulation and about 33 per cent when unmodulated.

The approximate treatment of Sec. 109 is applicable to Class B operation with fairly good accuracy. In this case, $E_a = 0$, and $\theta_1 = 90$ degrees, so that

$$E_b = \mu E_c \tag{46}$$

Equation (43) becomes

$$\begin{aligned} I_{p_1} &= \frac{E' g_m}{2} = \frac{\mu}{2r_p}\left(E_g - \frac{I_{p_1} R_b}{\mu}\right) \\ &= \frac{\mu E_g}{2r_p + R_b} \end{aligned} \tag{47}$$

From (47) it is evident that the apparent internal resistance of a tube operating Class B is twice the normal internal resistance r_p that it would have under Class A operation.

The direct-current component of plate current given by (44) becomes

$$I_b = \frac{E' g_m}{\pi} = \frac{2}{\pi} I_{p_1} = 0.637 I_{p_1} \tag{48}$$

The plate efficiency, from (35), is

$$\text{Plate efficiency} = \frac{E_o I_{p_1}}{2 E_b I_b} = \frac{\pi}{4} \frac{E_o}{E_b} \tag{49}$$

Since the maximum amplitude of the alternating voltage E_o across the tank circuit approaches E_b as a limiting value, as will be seen from Fig. 177, it follows from (49) that the plate efficiency of a Class B amplifier approaches 78.54 per cent as a limit.

112. Class C Amplifiers.—The foregoing treatment based upon the approximate characteristic may be also applied to Class C amplifiers, but with usually somewhat less accuracy because of the greater values of grid-excitation voltage commonly employed. With large values of excitation the upper part of the characteristic deviates appreciably from a straight line owing to saturation effects caused by either excessive grid current or insufficient emission from the filament.

Equation (43) now becomes

$$I_{p_1} = \frac{g_m E'}{\pi}\left(\theta_1 - \frac{1}{2}\sin 2\theta_1\right) = \frac{\mu}{\pi r_p}\left(E_g - \frac{I_{p_1}R_b}{\mu}\right)\left(\theta_1 - \frac{1}{2}\sin 2\theta_1\right)$$
$$= \frac{\mu E_g}{\dfrac{\pi}{\theta_1 - \frac{1}{2}\sin 2\theta_1}r_p + R_b} = \frac{\mu E_g}{\beta r_p + R_b} \tag{50}$$

where

$$\beta = \frac{\pi}{\theta_1 - \frac{1}{2}\sin 2\theta_1} = \frac{\pi}{\theta_1 - \sin\theta_1\cos\theta_1} \tag{51}$$

The term βr_p is the apparent internal resistance of the tube under Class C operation and is a function of the portion of the cycle during which plate current is allowed to flow. The value of β will be 2 for Class B operation when $\theta_1 = 90$ degrees, increasing to 34.68 for $\theta_1 = 30$ degrees.

Class C amplifiers are used in radio-telegraph transmitters and in radio telephony in the stages preceding the modulated stage. They cannot ordinarily be used to amplify a modulated signal since they are biased beyond cut-off and would distort the modulated envelope unless the percentage of modulation were small. This is illustrated in Fig. 185. Amplifiers of this type have been occasionally used to increase the percentage of modulation when the excitation voltage impressed on the amplifier is insufficiently modulated. This can be accomplished by choosing E_c so that the minimum positive amplitudes of the excitation voltage E_g at point a in Fig. 185 do not carry the potential of the grid below cut-off. In this way it is theoretically possible for a feebly modulated input to emerge completely modulated. Practically, the increase in the percentage of modulation that can be secured in this manner is limited, as the high value of bias needed would tend to reduce the linearity of the

dynamic characteristic, with consequent distortion of the audio-frequency envelope.

In the older types of radio-telephone transmitters the oscillator tube was modulated. This was accomplished by super-

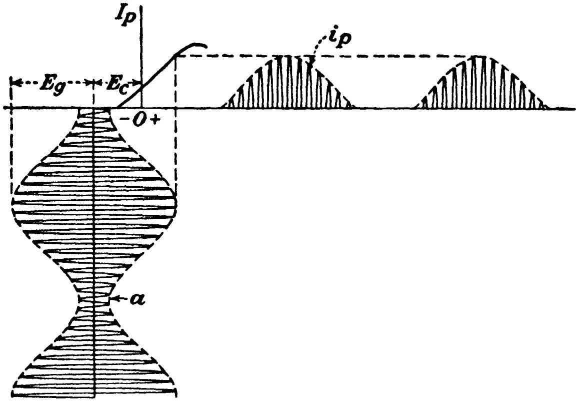


FIG. 185.—Distortion of the modulated envelope by a Class C amplifier.

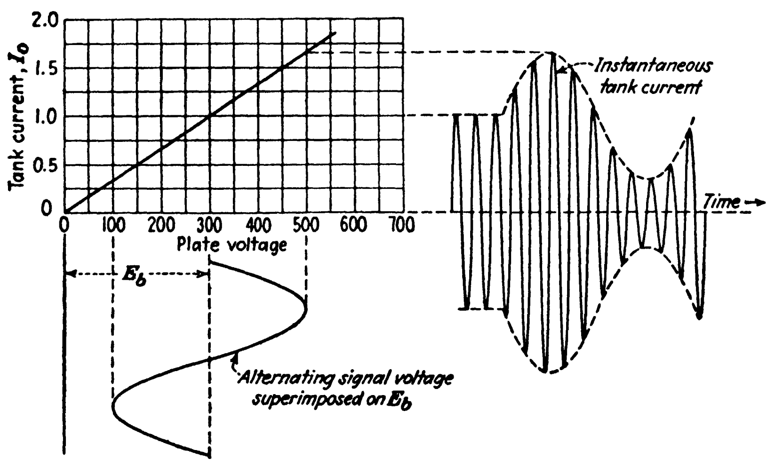


FIG. 186.—Modulation of a Class C amplifier by superimposing an audio-frequency signal voltage on the plate-supply voltage.

imposing an audio-frequency voltage upon the direct-current plate voltage. The average plate-supply voltage therefore rose and fell at an audio-frequency rate, causing the radio-frequency tank current to do likewise. This fluctuating plate voltage unfortunately caused the carrier frequency to flutter, which was objectionable. Present practice is to use a crystal-controlled

oscillator of relatively low power and amplify its output by successive stages of amplification. One of the intervening amplifier stages is then modulated. This modulated stage is operated Class C. The magnitude of the impedance offered by the tank circuit, the grid-excitation voltage, and the C bias, must all be chosen so that a linear relation exists between the tank current I_o and the voltage applied to the plate. Then, as the plate voltage rises and falls in accordance with the modulating frequency, the amplitude of the tank current will likewise rise and fall, as shown in Fig. 186. Any curvature in the characteristic will cause the modulated envelope to be distorted. The use of a grid leak and condenser for a portion of the C bias, in conjunction with a fixed voltage for the remainder, is sometimes of assistance in securing a linear characteristic.

113. Frequency Multipliers.—The plate current of a Class C amplifier is badly distorted and therefore contains a large percentage of harmonics. It is possible to resonate the tank circuit to one of these harmonics and cause it to absorb a large amount of power at this harmonic frequency. The impedance offered to the fundamental and the balance of the harmonics will be small and very little power will be absorbed at these other frequencies. The voltage across the tank circuit will now consist chiefly of this higher harmonic frequency.

Frequency multipliers are used to obtain higher frequencies than can be readily produced by a piezoelectric crystal-controlled master oscillator. The frequency of oscillation of the quartz crystal varies inversely with its thickness, so that a crystal ground to have very high natural mechanical period of vibration would be thin and very apt to crack in service. In order to secure crystal control of the frequency in the case of short-wave transmitters the crystal is ground to oscillate at some low-frequency multiple of the transmitted frequency. The output of the crystal-controlled oscillator is then impressed on one or more amplifiers connected in cascade and adjusted to multiply the frequency. The usual practice is to double the frequency with each stage, and while greater multiplications than this can be obtained, the plate efficiency, and hence the output, falls off rapidly as higher multiplications per stage are attempted. A Class C amplifier having a plate efficiency of 80 per cent would show an efficiency of about 70 per cent when used as a frequency doubler.

The instantaneous current and voltage relations of a frequency doubler are shown in Fig. 187. The plate voltage e_p will be low in value for a smaller time interval in this case than in Fig. 177, so that a smaller value of θ_1 must be used with frequency doublers in order to keep the losses within the tube small. These losses are proportional to the product of the instantaneous values of e_p and i_p and can be minimized by restricting the flow of plate current to a brief interval of time. This calls for values of E_g

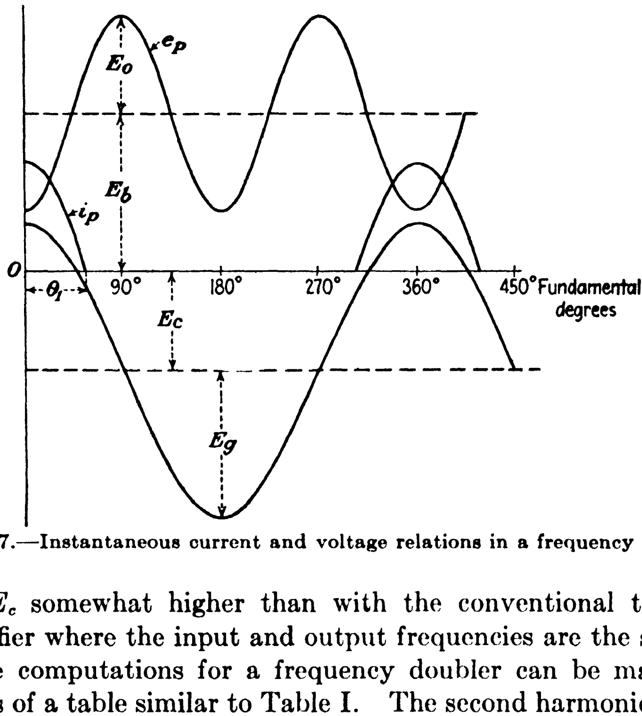


FIG. 187.—Instantaneous current and voltage relations in a frequency doubler.

and E_c somewhat higher than with the conventional type of amplifier where the input and output frequencies are the same.

The computations for a frequency doubler can be made by means of a table similar to Table I. The second harmonic I_{p_2} of the plate current now supplies the power to the tank circuit and may be evaluated by multiplying both sides of (21) by cos $2\theta d\theta$ and integrating the resulting equation between the limits zero and 2π. Solving for the amplitude of the second harmonic gives us

$$I_{p_2} = \frac{1}{\pi}\int_0^{2\pi} i_p \cos 2\theta d\theta = \frac{2}{\pi}\int_0^{\pi} i_p \cos 2\theta d\theta \qquad (52)$$

The definite integral can be evaluated by means of the trapezoidal

rule, as before. The balance of the computations will be similar to those for the output at fundamental frequency.

Either triodes or tetrodes can be used as frequency multipliers. The former will not need to be neutralized, as the input and output circuits are tuned to different frequencies and hence will not oscillate.

114. Neutralization of Power Amplifiers.—Radio-frequency power amplifiers using triodes exhibit the same tendencies to oscillate as do those discussed in the preceding chapter, and must

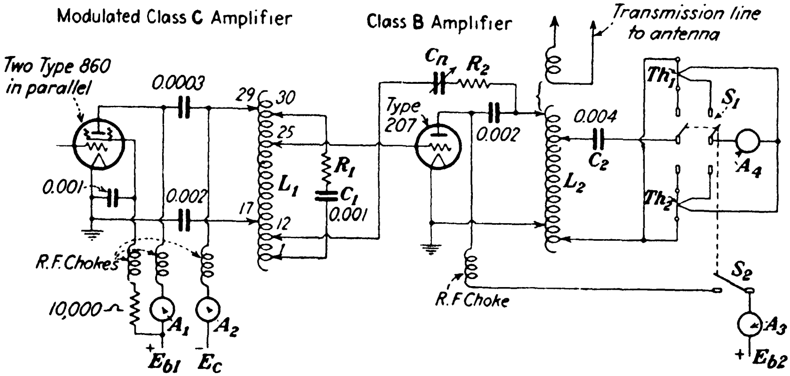


FIG. 188.—Modulated stage supplying grid excitation for the neutralized power amplifier of a typical broadcast transmitter.

therefore be neutralized. Any of the circuits discussed in Sec. 101 can be used. The problem of securing stability over a wide range of frequencies is not a factor since the tuning adjustments in transmitters remain fixed.

The adjustment of the neutralizing condenser to the proper value is generally made by means of a suitable ammeter in the output tank circuit which serves as the galvanometer in the equivalent bridge circuit. Figure 188 shows the last two stages of power amplification of a typical 1000-watt broadcast transmitter. The first stage is the modulated amplifier consisting of two 75-watt screen-grid tubes in parallel which require no neutralization. The second stage operates Class B and is neutralized by means of the condenser C_n which is connected from the plate lead to the input tank circuit L_1C_1 at the point shown. The principle is the same as that of the Rice circuit of Fig. 166. The turns to which the various taps on L_1 are connected are

indicated by the numbers. A resistance R_2 of about 3C ohms is connected in series with C_n to secure a more exact phase balance, since the bridge arms are not all pure reactances.

The neutralizing adjustment is made as follows: The switch S_1 is thrown to the top position which inserts a low-range thermocouple Th_1 in the output tank circuit L_2C_2. At the same time the galvanometer A_4 is connected to the thermocouple and the plate circuit is opened by S_2, which is mechanically connected with S_1. Excitation is applied to the grid and the balancing condenser is adjusted until A_4 reads zero. The switch S_1 is then thrown to the lower position, closing the plate circuit and inserting a high-range thermocouple Th_2 in the tank circuit, and at the same time transferring A_4 from Th_1 to Th_2.

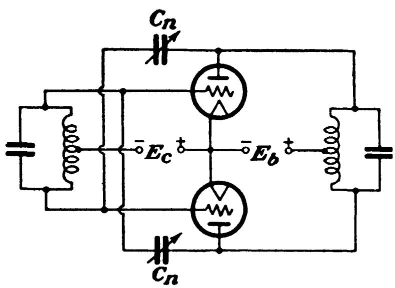


FIG. 189.—Neutralizing circuit for a push-pull amplifier.

The input tank circuit contains a resistance R_1 which increases the circuit losses so that the grid-excitation power taken by the second tube is not too great a percentage of the total output of the preceding stage. It also reduces the value of Q for this circuit. Too high a value of Q would be objectionable in the case of a modulated signal as the increased sharpness of the resonance curve would tend to discriminate against the outer side-band frequencies, resulting in frequency distortion.

Figure 189 shows in schematic form the method of neutralizing a push-pull amplifier. The arrangement is seen to be merely an application of the Rice circuit to each tube.

115. Crystal Oscillators.[5]—In order to secure the high degree of frequency stability required by modern systems of radio communication the tuned circuit of the oscillator is replaced by a piezoelectric quartz crystal, the resonating properties of which were discovered by Dr. W. G. Cady.

When a crystal possessing piezoelectric properties is subjected to a mechanical stress in a particular direction, electrical charges

[5] For an extensive bibliography on piezoelectricity see *Proc. I.R.E.*, vol. 16, p. 521, April, 1928.

will be produced. If the stress is changed from compression to tension, the polarity of the charges will reverse. The phenomenon is a reversible one, so that if a difference of potential is applied to the crystal, a mechanical deformation results. This effect was discovered by the Curies and is exhibited by a number of crystalline substances. It is most pronounced in Rochelle salts, and a section cut from a large crystal of this material is capable of producing a difference of potential sufficient to jump across a small spark gap connected to metallic-foil surfaces

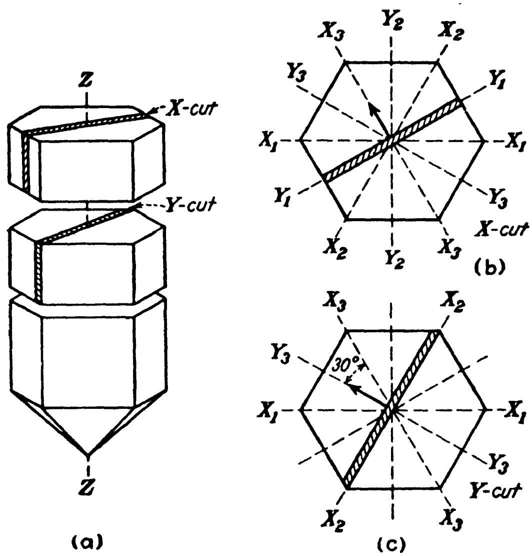


FIG. 190.—Method of securing X-cut and Y-cut quartz plates.

cemented to opposite faces of the crystal when the latter is rigidly clamped and struck a sharp blow with a lead pencil.

Quartz, while not possessing piezoelectric properties to such a marked degree, is more satisfactory in an oscillator as it is not affected by moisture and has a very low temperature coefficient. A perfect quartz crystal would have the appearance of Fig. 190*a*. The longitudinal axis Z passing vertically through the center of the crystal is called the optical axis. The three axes X_1, X_2, and X_3 passing through the corners of the hexagon are the electrical axes; while those perpendicular to the crystal faces, marked Y_1, Y_2, and Y_3, are the mechanical axes.

If a flat section is cut from the crystal so that the flat sides are perpendicular to any electrical axis such as X_3, as shown in Fig.

190*b*, mechanical stresses along the Y_1 axis will produce electrical charges along the flat sides of the slab. The section cut in this manner is called "*X* cut" or "Curie cut." If the plate is inserted between two fixed metal electrodes so as to become the dielectric of a small condenser, as illustrated in Fig. 191, the crystal will be set into mechanical vibration when an alternating voltage is applied. When the impressed frequency coincides with the natural mechanical period of vibration of the crystal the amplitude of vibration is very greatly increased, and if the voltage impressed is sufficiently large, the crystal will be shattered.

A section cut in the manner of Fig. 190*c* is called a *Y* or 30° cut and will usually oscillate more readily than the *X*-cut plates, particularly for the lower frequencies.

There will usually be several frequencies in a plate of rectangular shape corresponding to the various possible modes of mechanical vibration.[6] For the higher frequencies the thickness is the governing dimension so that the plate must be accurately ground to secure the desired frequency. The frequency of oscillation is inversely proportional to the thickness, a 1-mm. *Y*-cut plate having a resonant frequency of about 2×10^6 cycles. An *X*-cut plate of the same thickness would have a resonant frequency of approximately 3×10^6 cycles.

The *Y*-cut plates have a positive temperature coefficient, the natural frequency rising with increasing temperature. The increase in frequency varies from about 25 to 100 parts in a million per degree centigrade. *X*-cut crystals have a negative temperature coefficient, causing the natural frequency to decrease as the temperature rises. The amount in this case is about 10 to 25 parts in a million per degree centigrade. It has been found possible to make these two temperature effects approximately neutralize one another by cutting a ring-shaped piece from a *Y*-cut slab.[7]

Present regulations governing broadcast stations in the United States require that the carrier frequency must not deviate by more than 50 cycles from the assigned value. This requires an

[6] F. R. Lack, Observations on Modes of Vibration and Temperature Coefficients of Quartz Crystal Plates, *Proc. I.R.E.*, vol. 17, p. 1123, July, 1929.

[7] W. A. Marrison, A High Precision Standard of Frequency, *Proc. I.R.E.*, vol. 17, p. 1103, July, 1929.

accurate control of the crystal temperature, which is accomplished by placing the crystal in an electrically heated oven, the interior temperature of which is maintained constant to within a small fraction of a degree by a suitable thermostat. This enables the crystal to be ground approximately to the desired frequency, the exact value being then secured by raising or lowering its operating temperature. The length of the air gap between the crystal and the upper plate of the holder in Fig. 191 also has an appreciable effect upon the frequency of oscillation so that a crystal should always be calibrated in its holder.[8] The latter should be designed so that the length of the gap is accurately maintained. Supersonic air waves are produced in the air within the gap, and if the length of the gap is an integral number of half wave lengths, considerable energy may be absorbed by the air column. This will place an appreciable load on the crystal and may even prevent it from oscillating. The length of the air gap must therefore be selected so as to avoid this possibility.

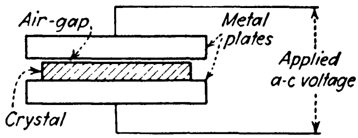


FIG. 191.—Method of mounting a piezoelectric crystal.

The equivalent electrical network of the vibrating crystal resonator is shown in Fig. 192.[9] The condenser C_1 represents the capacitance between the two plates of the holder of Fig. 191 when the crystal is not vibrating. The series-resonant circuit, L, C, and R, represents the electrical equivalent of mass, compliance (the reciprocal of stiffness), and friction. The resonant frequency of L and C is the frequency of mechanical resonance. The energy consumed by this equivalent circuit from the source of E will be relatively large at resonance and will be very small for higher or lower frequencies. At frequencies above resonance the inertia of the crystal causes the resulting vibration to lag behind the applied force so that energy is supplied at a lagging power factor. Under these circumstances

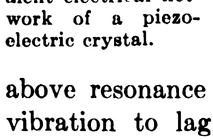


FIG. 192.—Equivalent electrical network of a piezoelectric crystal.

[8] A. HUND, Notes on Quartz Plates, Air-gap Effect, and Audio-frequency Generation, *Proc. I.R.E.*, vol. 16, p. 1072, August, 1928.

[9] K. S. VANDYKE, The Piezo-electric Resonator and Its Equivalent Network, *Proc. I.R.E.*, vol. 16, p. 742, June, 1928.

the crystal will possess inductive reactance, although the structure consists of two metal plates with a quartz dielectric between them.

The effective value of Q of the equivalent circuit is much higher than that of the ordinary coil and condenser, particularly when the resonant frequency of the crystal is low. Values of Q will vary from about 1000 for a 1000-kc crystal to perhaps 10,000 for one ground for 100 kc. The value of C in Fig. 292 is only a few hundredths of a micro-microfarad, while L is large, varying from a fraction of a henry to 100 henrys, or more, for low-frequency crystals. Quartz crystals have been used in electrical filters in cases where very sharply defined cut-off frequencies are desired.[10] The high value of Q possessed by a vibrating crystal makes it well suited for such applications.

The use of a crystal for controlling the frequency of a vacuum-tube oscillator is illustrated in Fig. 193. The two plates of the holder are connected between the grid-filament terminals of a small power tube having a parallel-resonant circuit in the plate circuit. When the latter is adjusted to approximate resonance, the resistance component R_g of the input impedance becomes sufficiently negative to annul all of the positive resistance of the equivalent crystal circuit, and oscillations begin. The energy to maintain these oscillations is fed back through the grid-plate capacitance of the tube. The frequency is dictated by the mechanical period of vibration of the crystal and the plate circuit may be considerably detuned without materially affecting the frequency. A negative bias, if needed, may be applied either through a grid leak R_c as shown, or through a choke coil. With small tubes operating at moderate values of plate voltage the grid bias may be omitted. A 50-watt tube is about the largest size that can be safely used with a crystal, particularly at higher frequencies where the crystal is thin and fragile.

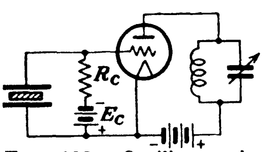


FIG. 193.—Oscillator circuit for crystal control.

The frequency range of crystal oscillators varies from about 30 kc to about 4000 kc as a practical upper limit. Where higher frequencies are desired it is necessary to use a frequency multiplier.

[10] W. P. MASON, Electrical Wave Filters Employing Quartz Crystals as Elements, *Bell System Tech. Jour.*, vol. 13, p. 405, July, 1934.

116. Magnetostriction Oscillators.—Another form of oscillator circuit wherein the frequency is controlled by mechanical resonance is the magnetostriction oscillator due to G. W. Pierce.[11] These oscillators are well suited for applications in the lower range of frequencies where crystal control is impractical because of the difficulty in securing quartz crystals of sufficient size. The useful frequency range extends from 300 kc to perhaps as low as 500 cycles.

The phenomenon of magnetostriction relates to the stresses and changes in dimensions produced in a material by magnetization, and the inverse effect of changes in the magnetic properties produced by mechanical stresses. Nickel, alloys of nickel and iron, invar, nichrome, and various other alloys of iron exhibit pronounced magnetostrictive effects. Pure iron and iron with various amounts of carbon are unsatisfactory.

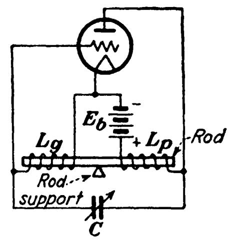


FIG. 194.—Magnetostriction oscillator circuit.

The circuit of a magnetostriction oscillator is shown in Fig. 194. The arrangement is similar to a Hartley circuit, except for the reversed mutual inductance between the grid and plate coils. This prevents the circuit from oscillating, except when the magnetostrictive rod is inserted. The rod is pivoted or clamped at the center and is magnetized by the direct-current component of the plate current. When oscillations are started, the rod will vibrate longitudinally with a node at the center, *i.e.*, it tends to lengthen and shorten at a frequency given by

$$f = \frac{v}{l} \tag{53}$$

where v is the velocity of sound in the rod and l is the length of the rod.

As in the case of the crystal oscillator, the condenser C can be varied over a considerable range without appreciably changing the frequency.

117. Frequency Stability of Oscillators.—It is often impractical to employ mechanical resonance as a means of controlling the frequency of an oscillator. This is particularly true in the case of oscillators of adjustable frequency for laboratory purposes.

[11] *Proc. I.R.E.*, Magnetostriction Oscillators, vol. 17, p. 42, January, 1929.

As pointed out in Sec. 104, the frequency of oscillation is affected to some extent by the constants of tube used and by the resistance of the load. This relation, from (9), is

$$f = f_o\sqrt{1 + \frac{R_o}{r_p}}$$

Changes in the frequency are due chiefly to variations in the filament and plate-supply voltages which cause corresponding variations in r_p. These changes can be minimized by means of the circuit given in Fig. 195*a*.[12] The output voltage of the tube is impressed across two resistances R_1 and R_2 in series, R_1 being made large compared to R_2 so as to reduce the effects of changes in the load impedance. The parallel-resonant circuit L_oC_o is connected to the output of the tube through a very high feed-back

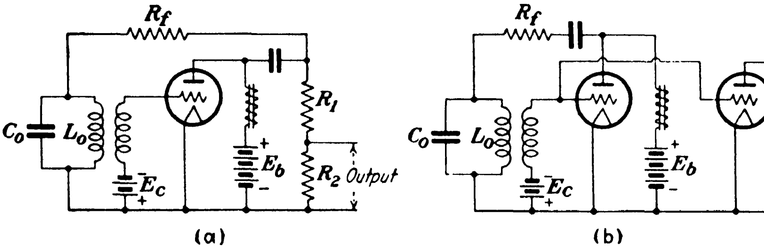


FIG. 195.—Oscillator circuits having high frequency stability.

resistance R_f. This resistance is in series with r_p of the tube and serves to reduce the effects of any changes in the latter.

This arrangement materially reduces the output and efficiency of the oscillator so that subsequent amplification of the output is usually required. In this case the input terminals of the amplifier tube can be connected across R_2 which may be made variable to serve as an output control. An alternate scheme, shown in Fig. 195*b*, does away with R_1 and R_2 and connects the grid of the amplifier tube directly to the oscillator grid.

Circuits of this type will not experience a change in frequency of more than a few hundredths of 1 per cent for relatively large variations in the plate and filament voltages. Frequency stabilization can also be secured by introducing a phase shift between the input and output voltages of the oscillator, provided the value of Q for the tuned circuit is very high, as pointed out

[12] J. W. HORTON, *Bell System Tech. Jour.*, vol. 3, p. 521, July, 1924.

by F. B. Llewellyn.[13] In several cases this phase shift can be obtained by a proper choice of the grid condenser in self-biased oscillators, or a similar choice in value of the plate blocking condenser.

These methods are unable to compensate for slight changes in the values of inductance and capacitance of the tuned circuit, such as those produced by variations in their temperature.

118. Beat-frequency Oscillators.—In the ordinary type of audio-frequency oscillator the frequency is adjusted by changing the constants of the tuned circuit by using a tapped coil in con-

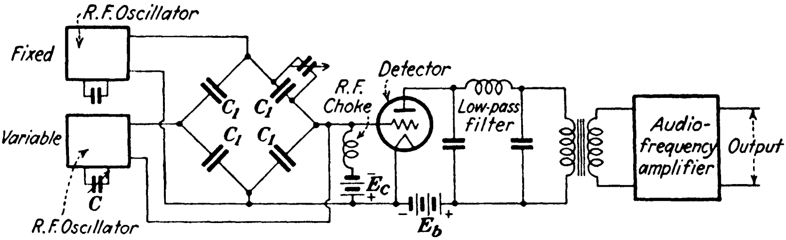


FIG. 196.—Schematic circuit of a beat-frequency oscillator using a balanced capacitance bridge to avoid coupling between the two radio-frequency oscillators

junction with a decade condenser shunted by a variable air condenser. This requires the manipulation of several dials so that the frequency cannot be varied continuously over the entire tuning range without interruption. This may be objectionable, particularly in connection with acoustic tests on loud-speakers where it is often desirable to have a smooth frequency variation in order conveniently to detect resonant peaks in the output of the device.

A continuously variable frequency is readily obtained in a beat-frequency oscillator which utilizes the beats between the outputs of two radio-frequency oscillators having slightly different frequencies. The two radio frequencies are impressed on a suitable vacuum-tube detector which rectifies the impressed voltage and extracts the beat frequency. This process is called *heterodyning*. The detector output is passed through a suitable low-pass filter and into an audio-frequency amplifier, as shown in Fig. 196. Thus, one oscillator frequency may be fixed at 100,000 cycles and the other made adjustable from 90,000 to 100,000

[13] Constant Frequency Oscillators, *Proc. I.R.E.*, vol. 19, p. 2063, December, 1931; or *Bell System Tech. Jour.*, vol. 11, p. 67, January, 1932.

cycles by means of a single variable condenser giving an audio-frequency range extending from zero to 10,000 cycles. Provision must be made to avoid coupling between the two oscillators as they will tend to synchronize automatically when the difference between their two frequencies is small. This must be prevented if low values of audio frequency are to be produced. Connecting the two oscillators to two pairs of diagonally opposite points of a balanced bridge circuit composed of four equal condensers C_1 is one way of avoiding this difficulty. The small variable condenser across one of the arms compensates for the input capacitance of

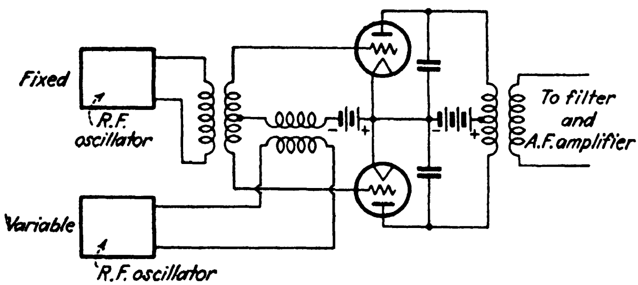


FIG. 197.—Beat-frequency oscillator using balanced detectors.

the detector. Another arrangement is to use two balanced detectors in a sort of push-pull circuit, as illustrated in Fig. 197.

The output of the detector will contain harmonics of the impressed frequencies as well as their sums and differences, as shown by (4), page 147. Since we are interested only in the difference between the fundamentals of the two impressed frequencies, a low-pass filter of one or more sections is inserted in the output of the detector, which transmits the difference frequency and eliminates the terms of higher order.

A small change in the frequency of either oscillator will result in a relatively large percentage change in difference frequency, so that the frequency stability of a beat-frequency oscillator is relatively poor. Accordingly the two radio-frequency oscillators should be designed to secure a high degree of frequency stability. Their circuit details should be as near alike as possible so that temperature changes, etc., will affect each oscillator to the same degree.

Another advantage of the beat-frequency oscillator is that its output voltage remains fairly constant over the entire frequency

range, assuming the audio-frequency amplifier to be reasonably free from frequency distortion. This is due to the relatively small variation in the frequency of the adjustable oscillator. It also enables very low frequencies to be obtained without the use of very large values of either inductance or capacitance that would be required by the ordinary type of oscillator circuit.

119. The Multivibrator.—Another type of oscillator circuit, known as the multivibrator, which also permits a fairly wide frequency range to be secured by means of a single adjustment, is

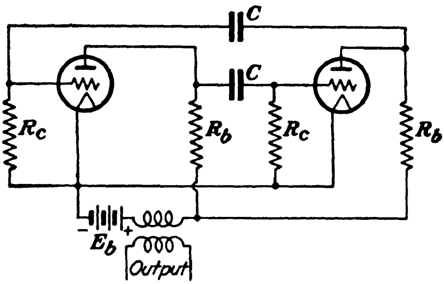


FIG. 198.—Circuit diagram of a multivibrator.

shown in Fig. 198. It is essentially a two-stage resistance-coupled amplifier with the output of the second stage fed into the grid of the first stage. The input and output voltages of each stage are 180 degrees out of phase with each other so that two stages are needed in order to secure the proper phase relations to support oscillations. The action of the circuit is similar in all respects to motor-boating in a resistance-coupled amplifier, discussed in Sec. 84.

The frequency of oscillation, since there is no tuned circuit present, is governed primarily by the time constant of the grid-leak and condenser R_cC and to a much lesser extent by the other circuit constants. The usual manner of varying the frequency is to vary the value of either or both blocking condensers. These may be large variable condensers or fixed condensers shunted by smaller variable condensers. Varying the value of R_c of either grid leak will produce the same results. The wave shape is badly distorted which seriously restricts its use as an oscillator.

Multivibrators are used chiefly as a source of harmonics in the measurement of frequency. Harmonics in the output as high as the 80th can be detected without the use of amplifiers. Thus,

if the fundamental frequency is accurately known by comparing it with the fundamental or one of the octaves of a standard tuning fork, a series of harmonic frequencies is obtained from the multivibrator which extends into the lower range of radio frequencies. The frequency range can be extended by synchronizing the fundamental of a second multivibrator with some higher harmonic of the first. Since the only inductance present is that of the various lead wires, the electrical inertia of the system is very small. This condition enables a standard frequency to be injected into the circuit in series with the plate-supply voltage which will coincide with a harmonic of the multivibrator and cause the latter to lock into synchronism with the injected frequency.[14] The fundamental of the multivibrator then becomes an integral submultiple of the injected control frequency, thus making the device a step-down frequency converter. In this way a control frequency derived from a crystal oscillator may be reduced to a value low enough to drive a suitable counting mechanism, such as a motor-driven clock. The frequency of the crystal oscillator can be established to a high order of accuracy in this way by observing the errors in time as indicated by the clock.

120. Dynatron Oscillators.—The effects of secondary emission have already been mentioned in connection with Fig. 181 in the case of a triode and in Fig. 135 in the case of a tetrode. When the grid of a triode is operated at a higher positive potential than that of the plate, a characteristic may be obtained similar to Fig. 199. In the range of plate voltages between b and d the number of secondary electrons leaving the plate exceeds the number of primary electrons arriving and the plate current reverses. Whether or not an actual reversal is obtained will depend upon surface conditions of the plate, electrode voltages, etc. The plate resistance of the tube is negative within the interval between a and c, and if a tuned circuit is inserted in the plate lead, as in Fig. 200a, the negative value of plate resistance may be sufficient to annul the positive resistance of the tuned circuit so that oscillations will occur. A tube operated in this manner was first described by A. W. Hull who gave it the name "dynatron."[15]

[14] L. M. HULL and J. K. CLAPP, A Convenient Method for Referring Secondary Frequency Standards to a Standard Time Interval, *Proc. I.R.E.*, vol. 17, p. 252, February, 1929.

[15] *Proc. I.R.E.*, vol. 6, p. 5, February, 1918.

A screen-grid tube is usually more satisfactory as a dynatron oscillator than a triode as the frequency stability is greater. In Fig. 135, the static I_p-E_p characteristics are suitable for dynatron operation in the region to the left of line A where the plate voltage is lower than the screen-grid voltage. The control

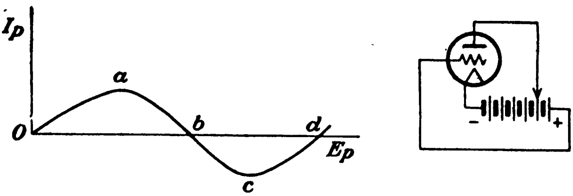


FIG. 199.—Dynatron characteristic of a triode obtained when the grid is at a higher positive potential than the plate.

grid is usually connected directly to the cathode, although a small negative bias will increase the negative value of r_p. The circuit diagram is shown in Fig. 200*b*.

The frequency stability of a dynatron using an ordinary screen-grid tube is somewhat comparable to a crystal oscillator without temperature control. This enables them to be used as heterodyne wave meters, the wave meter serving as the tuned circuit. If a

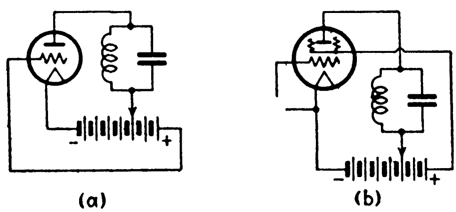


FIG. 200.—Dynatron oscillator circuits.

pair of telephone receivers are inserted in the screen-grid circuit of the oscillating dynatron, beats can be heard between the dynatron and other oscillators at harmonics of the dynatron frequency as well as at the fundamental. This enables the range of the wave meter to be extended into the higher frequencies.

The dynatron can also be used to measure the high-frequency resistance of a tuned circuit.[16] The negative resistance of the tube is varied by means of an adjustable bias impressed on the control grid until oscillations just begin. The negative resistance

[16] H. IINUMA, A Method of Measuring the Radio-frequency Resistance of an Oscillatory Circuit, *Proc. I.R.E.*, vol. 18, p. 537, March, 1930.

of the tube is then equal to the impedance of the parallel-resonant circuit and will be

$$r_p = \frac{L}{CR} \tag{54}$$

where C, L, and R are the constants of the tuned circuit.

121. Oscillating Arc.—Any device possessing a negative resistance may be used to produce oscillations by the proper introduction of inductance and capacitance into the circuit. The characteristic of the electric arc shown in Fig. 201 has a negative slope, indicating that the internal resistance is negative, so that it is capable of producing oscillations. The arc converter, or Poulsen arc, was developed by Valdemar Poulsen of Denmark and was very widely used as a source of high-frequency oscillations in radio transmitters prior to the introduction of the vacuum tube in this field.

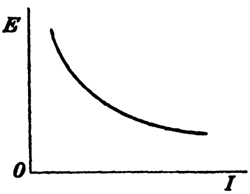


FIG. 201.—Characteristic curve of an electric arc.

The negative-resistance characteristic of the arc can be greatly accentuated by enclosing it in a hydrocarbon atmosphere and passing a strong transverse magnetic field across the arc stream.[17] A further improvement results if the positive electrode is made of copper and is water-cooled. The schematic diagram of connections is shown in Fig. 202. Shunted across the arc is the oscillatory circuit L_1C_1, which is of the series-resonant type in this case, as the internal resistance of the source is relatively low. This circuit is made up of the constants of the antenna with sufficient loading inductance connected in series for tuning purposes. The magnetic field across the arc is produced by coil L_2, while L_3 is a radio-frequency choke coil to exclude radio-frequency current from the direct-current generator. Signaling is accom-

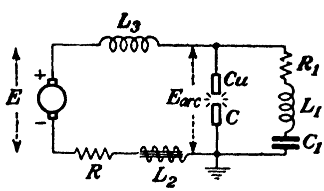


FIG. 202.—Schematic diagram of an arc converter.

[17] P. O. PEDERSEN, On the Poulsen Arc and Its Theory, *Proc. I.R.E.*, vol. 5, p. 255, August, 1917; also L. F. FULLER, The Design of Poulsen Arc Converters for Radio-telegraphy, *Proc. I.R.E.*, vol. 7, p. 449, October, 1919.

plished by short-circuiting a portion of the inductance in series with the antenna which changes the frequency of oscillation slightly when the key is depressed. Two different frequencies were therefore radiated—a signaling frequency and a spacing frequency. A heterodyne system of reception was used.

Arc converters were built having ratings of several thousand kilowatts. They have practically all been replaced by vacuum-tube transmitters operating at short wave lengths, giving the same transmission range at a fraction of the power required by the arc. The latter was primarily a low-frequency device and unsatisfactory for frequencies much above 100 kc. The radiated

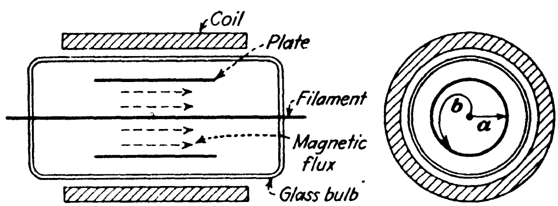


FIG. 203.—Construction of a magnetron showing electron paths.

wave was very rich in harmonics which produced considerable local radio interference.

122. Magnetron Oscillators.—The magnetron is an evacuated tube of the diode type having a cylindrical anode, along the axis of which is located the filament. Instead of a grid, the flow of electrons to the positively charged anode is controlled by an axial field produced by a coil wound outside the tube, as shown in Fig. 203.[18] In the absence of any magnetic field the electrons will travel from the filament to the plate in radial lines, as in *a*. Since these electrons in motion constitute a current, an axial magnetic field parallel to the filament will produce a tangential force on the moving electron causing it to follow a spiral path as in *b*. A field of sufficient strength will cause the radius of curvature of the path to be just equal to the radius of the plate and any further increase in the field strength above this critical value will prevent the electrons from ever reaching the plate. Flux densities below the critical value will have practically no effect upon the plate current as the number of electrons reaching

[18] A. W. HULL, The Magnetron, *Jour. A.I.E.E.*, vol. 40, p. 715, September, 1921; also The Axially Controlled Magnetron, *Trans. A.I.E.E.*, vol. 42, p. 915, 1923.

the plate will be the same, the only difference being in the length of the spiral path.

The characteristics of a typical magnetron showing the variation in plate current with the magnetic field for various values of plate voltage are illustrated in Fig. 204. The plate in this case was a cylinder 2 in. in diameter and 2 in. long.

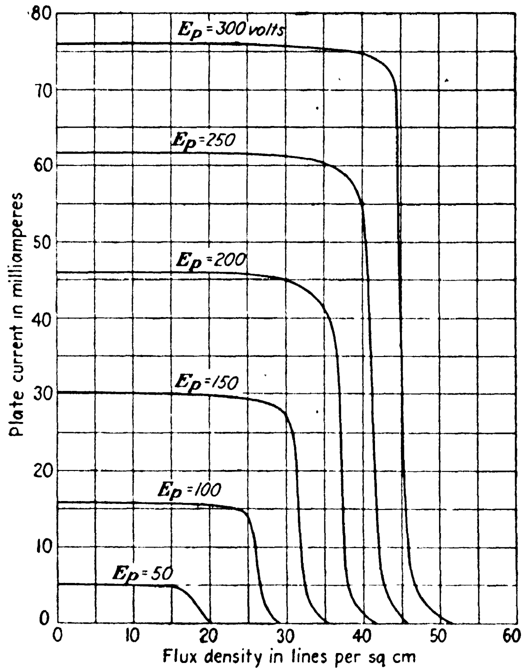


FIG. 204.—Effect of magnetic field on plate current of a magnetron at various plate voltages.

The magnetron may be used as either an amplifier or an oscillator.[19] In the latter capacity it has no particular advantages over the triode, except at very high frequencies.[20] When used as a generator of very high frequency oscillations the cylindrical plate is split into two halves and connected to a tuned circuit, as shown in Fig. 205. The electrons flowing from filament to

[19] F. R. ELDER, The Magnetron Amplifier and Power Oscillator, *Proc. I.R.E.*, vol. 13, p. 159, April, 1925.

[20] W. C. WHITE, Producing Very High Frequencies by Means of the Magnetron, *Electronics*, vol. 1, p. 34, April, 1930; also E. D. MCARTHUR and E. E. SPITZER, Vacuum Tubes as High-frequency Oscillators, *Proc. I.R.E.*, vol. 19, p. 1971, November, 1931.

plate under the combined influence of the electromagnetic and electrostatic fields give the volt-ampere characteristic of the tube a negative slope in the vicinity of cut-off, which can be utilized to produce oscillations in a tuned circuit in much the same manner as the dynatron. Each anode delivers energy to the tank circuit during alternate half cycles, so that the circuit is equivalent to a push-pull oscillator. The direct-current magnetic field produced by the coil is adjusted to a value slightly greater than cut-off.

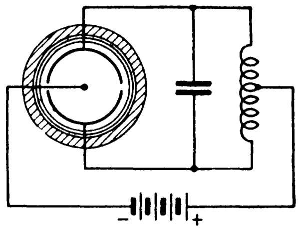


FIG. 205.—Split-anode type of magnetron oscillator used to generate very high frequencies.

123. Barkhausen Oscillations.[21]—When attempts are made to generate very high frequencies having wave lengths in the vicinity of one meter by conventional feed-back methods, pronounced difficulties occur in attempting to secure the proper phase relations between the alternating grid and plate potentials. This is caused by the period of oscillation becoming comparable to the time of transit of the electrons. The fictitious voltage μe_g acting in the plate circuit will therefore lag behind the alternating grid potential so that μ must be given a phase angle and treated as a complex number at these very high frequencies.

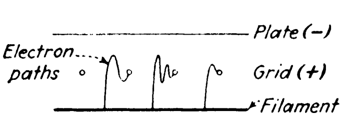


FIG. 206.—Paths of electron oscillations in a Barkhausen oscillator.

The high-frequency limitations of the ordinary type of oscillator may be avoided by utilizing the oscillations of the electrons. If the grid of a triode is maintained at a high positive potential with respect to the filament while the plate is kept at a much smaller negative potential with respect to the filament, electrons will be attracted to the grid with a high velocity. Those which do not strike the grid wires will make one or more oscillations before being finally drawn to the grid, as illustrated in Fig. 206.

[21] H. BARKHAUSEN and K. KURZ, Shortest Waves Obtainable with Valve Generators, *Ztschr. Phys.*, vol. 21, p. 1, January, 1920. A comprehensive discussion of electron oscillations is given by H. E. HOLLMANN, On the Mechanism of Electron Oscillations in a Triode, *Proc. I.R.E.*, vol. 17, p. 229, February, 1929; see also E. C. MEGAW, Electronic Oscillations, *Jour. I.E.E.* (*London*), vol. 72, p. 313, April, 1933.

The frequency of these oscillations is governed by the geometry of the tube and the potential of the grid, and not by the constants of the external circuit. A tube with a cylindrical grid and plate is necessary, although a special type of tube employing the conventional "flat-type" construction with grid and plate forming parallel planes has been successfully used.[22]

The wave length in centimeters to a rough degree of approximation is given by Barkhausen and Kurz as

$$\lambda = \frac{1000d}{\sqrt{E_g}}$$

where d is the diameter of the plate and E_g is the voltage of the grid.

The phase relations of these electron oscillations would appear to be entirely at random so that the net effect on the external circuit should be nil. However, the electrons surging back and forth in the grid and plate leads cause the potentials of these electrodes to rise and fall, which probably reacts on the random motion of the oscillating electrons and tends to coordinate their motion.

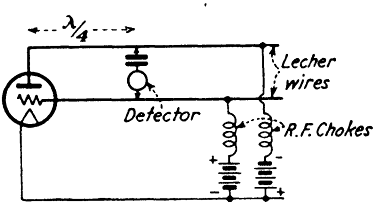


FIG. 207.—Circuit for generating Barkhausen oscillations.

Under certain circumstances the frequency of oscillation is found to depend upon the constants of the external circuit connected to the tube electrodes, particularly if the grid and plate terminate in a pair of Lecher wires as in Fig. 207.[23] These wires are merely a pair of parallel copper rods one-quarter wave in length. Their electrical properties are discussed in Sec. 195, Chap. XIV.

Oscillations of the Barkhausen type can also be produced in a magnetron, wave lengths of only a few centimeters having been obtained.[24]

[22] B. J. THOMPSON and P. D. ZOTTU, An Electron Oscillator with Plane Electrodes, *Proc. I.R.E.*, vol. 22, p. 1374, December, 1934.

[23] E. W. GILL and J. H. MORRELL, Short Electric Waves Obtainable by Valves, *Phil. Mag.*, vol. 44, p. 161, July, 1922.

[24] K. OKABE, On the Short Wave Limit of Magnetron Oscillations, *Proc. I.R.E.*, vol. 17, p. 652, April, 1929.

The energy obtained from these various types of electron oscillators is usually small, being ordinarily a fraction of a watt.

Problems

1. A type 801 triode having the characteristic curves given in Figs. 178 and 179 is used in the oscillator circuit of Fig. 174*a* under the following

CHARACTERISTICS OF TYPE 207 TRIODE*

$E_c = +300$		$E_c = 0$		$E_c = -300$		$E_c = -600$	
E_p	I_p	E_p	I_p	E_p	I_p	E_p	I_p
500	1.08	1,000	0.10	6,000	0.03	12,000	0.07
1000	1.65	2,000	0.26	7,000	0.15	13,000	0.18
1500	1.95	3,000	0.49	8,000	0.32	14,000	0.34
2000	2.21	4,000	0.79	9,000	0.54	15,000	0.55
3000	2.65	5,000	1.13	10,000	0.81	16,000	0.81
4000	3.04	6,000	1.51	11,000	1.13	17,000	1.09
$E_c = +200$		7,000	1.91	12,000	1.48	$E_c = -700$	
500	0.83	$E_c = -100$		$E_c = -400$		13,000	0.01
1000	1.05	3,000	0.10	8,000	0.04	14,000	0.08
1500	1.23	4,000	0.27	9,000	0.16	15,000	0.19
2000	1.42	5,000	0.50	10,000	0.32	16,000	0.35
2500	1.62	6,000	0.79	11,000	0.53	17,000	0.57
3000	1.82	7,000	1.13	12,000	0.79	18,000	0.83
4000	2.24	8,000	1.52	13,000	1.09	$E_c = -800$	
5000	2.67	9,000	1.93	14,000	1.45		
$E_c = +100$		$E_c = -200$		$E_c = -500$		15,000	0.01
						16,000	0.08
1000	0.47	5,000	0.11	10,000	0.06	17,000	0.21
1500	0.60	6,000	0.28	11,000	0.17	18,000	0.37
2000	0.75	7,000	0.50	12,000	0.33	19,000	0.60
3000	1.09	8,000	0.79	13,000	0.54	20,000	0.84
4000	1.45	9,000	1.12	14,000	0.80		
5000	1.83	10,000	1.50	15,000	1.11		
6000	2.23	11,000	1.91				

* I_p is in amperes.

operating conditions: $E_b = 500$ volts, $E_{pmin} = 125$ volts, $e_{gmax} = 100$ volts, $\theta_1 = 50°$, $C_o = 0.002\ \mu f$, $f = 10^6$ cycles. If $\mu = 8$, find E_o, E_g, I_b, I_c, I_{p1}, I_{g1}, power input from the B supply, power output to the tank circuit, plate efficiency, plate dissipation, net useful output, grid-leak resistance, M, L_o, and the effective resistance of the load R_o.

2. Repeat Problem 1 with θ_1 reduced to 40°, the other operating conditions remaining the same.

3. What would be the net output of the tube in Problem 2 when operated as a Class C power amplifier?

4. A type 207 triode, whose characteristics are given in the table, is operated as a Class B amplifier under the following conditions: $E_b = 15{,}000$ volts, $E_g = 1100$ volts, $E_c = -800$ volts, $\theta_1 = 90°$, $\mu = 20$, $E_{p\min} = 4000$ volts. From the data given in the accompanying table determine I_b, I_{p_1}, R_b, output, input, and the plate efficiency.

5. Using the approximate tube characteristics of Fig. 183, determine the same items as in Problem 4, assuming the mutual conductance of the 207 tube to be 5400 micromhos. Use the value of R_b as determined above.

6. The tube of Problem 1 is to be operated as a frequency doubler converting 2000 kc into 4000 kc, under the following operating conditions: $E_b = 500$ volts, $E_{p\min} = e_{g\max} = 100$ volts, $\theta_1 = 30°$. Find the power input from the B supply, the power output to the tank circuit, and the plate efficiency. If the tank condenser is 0.0005 μf, what will be the value of current flowing through it?

CHAPTER XI

MODULATION

124. Types of Modulation.—Modulation is defined as the process whereby the amplitude, frequency, or phase of a wave is varied in accordance with a signal wave. The resultant wave will contain frequencies identical with those present in the two original waves, and in addition, new frequencies will be produced which will be made up of various combinations of the original frequencies.

Assume a current wave defined by the expression

$$i = A \sin (\omega t + \phi) \tag{1}$$

where A is the amplitude, $\omega = 2\pi \times$ frequency, and ϕ is the

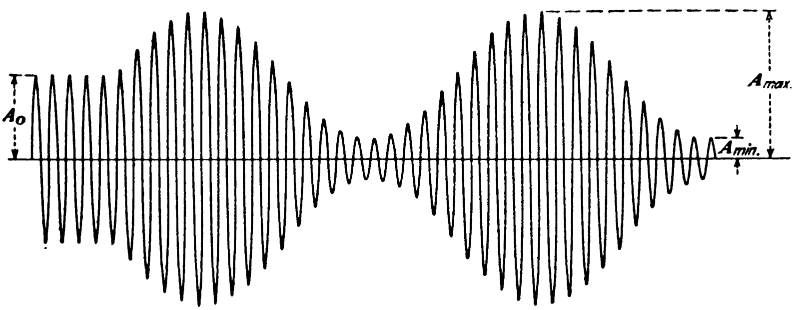


FIG. 208.—Amplitude-modulated wave.

phase angle. Any one of these three independent variables may be subjected to a periodic change which is slow compared to the "carrier" frequency ω; giving rise to amplitude, frequency, or phase modulation, respectively. A physical picture of these three types of modulation may be obtained by considering an ordinary type of alternator having the conventional revolving field. The frequency of the alternator will be governed by the speed and the number of poles. If the direct-current excitation of the field is varied sinusoidally by manipulating the field rheostat, the amplitude of the voltage induced in the stationary armature will rise and fall as illustrated in Fig. 208, assuming the

alternator to have a linear magnetization curve. This is in effect superimposing a very low frequency alternating current upon the direct-current field excitation. If a high-frequency alternator were used, it would be theoretically possible to introduce the amplified output of a microphone into the field circuit of the machine and thus modulate the voltage output of the alternator. This modulated wave when impressed upon a linear rectifier would result in a rectified output containing all of the original audio frequencies that were introduced into the field circuit of the alternator. Since this rectification or detection process is just the reverse of modulation, it is often called *demodulation.*

Frequency modulation would be accomplished in this example by causing the speed of the machine periodically to increase and decrease in accordance with the modulating frequency. Phase modulation could be brought about by rotating the stator of the alternator back and forth through an angle ϕ, assuming the stator to be mounted on suitable trunnion bearings so that it is free to rotate. Neither of these two latter types of modulation can produce an audio signal in the output of the rectifier. In order for a phase- or a frequency-modulated signal to become audible, it must be converted into an amplitude-modulated signal. This is accomplished by inserting selective tuned circuits ahead of the rectifier or else by heterodyning the modulated wave with a fixed frequency. Neither of these two types of modulation is suitable for radio telephony, except at extremely high frequencies where frequency modulation could perhaps be used. They are of interest chiefly in that either or both may be present as unwanted by-products in various systems of amplitude modulation. An example of simultaneous amplitude and frequency modulation is the case of an oscillator which is modulated by varying the plate-supply voltage at an audio-frequency rate. As pointed out in the preceding chapter, these variations in the voltage cause similar variations in the frequency of oscillation as well as in the amplitude of the tank current.

In continuous-wave radio telegraphy (abbreviated C.W.) the dots and dashes are produced by interrupting the carrier wave during the spacing intervals, which is a case of amplitude modulation. In the case of arc converters, "keying" was accomplished by means of frequency modulation, the signaling frequency being made slightly higher than the spacing frequency by short-

circuiting a portion of the tuning inductance when the key was depressed. The signals in both types of transmission are made audible by means of a heterodyne detector. This is necessary even in the case of the amplitude-modulated signals as otherwise the dots and dashes would be heard in the telephone receivers as a succession of almost unintelligible clicks. The radio-frequency oscillations are sometimes modulated at an audio-frequency rate at the transmitter by using 500-cycle alternating current as a source of plate voltage, or else by rapidly interrupting the oscillations by means of a motor-driven disk having insulated segments known as a "chopper." This type of signal is known as "interrupted continuous wave" (I.C.W.). An audible tone will be heard in the telephone receivers with this method when an ordinary rectifying detector is used.

125. Amplitude Modulation.—If the amplitude A in (1) is made to vary sinusoidally, then at any instant

$$A = A_0(1 + m \sin \omega_s t) \tag{2}$$

where m is the modulation factor and ω_s is the radian frequency of the modulating source. The factor m may have any value between zero and unity, and is defined, referring to Fig. 208, by

$$m = \frac{\frac{1}{2}(A_{max} - A_{min})}{A_0} \tag{3}$$

Substituting the value of A from (2) in (1), and assuming the phase angle ϕ in (1) to be zero, the equation of the wave is

$$i = A_0(1 + m \sin \omega_s t) \sin \omega_c t \tag{4}$$

where ω_c is the radian frequency of the carrier.

Expanding (4), we get

$$\begin{aligned} i &= A_0[\sin \omega_c t + m \sin \omega_s t \sin \omega_c t] \\ &= A_0\left[\sin \omega_c t + \frac{m}{2} \cos (\omega_c - \omega_s)t - \frac{m}{2} \cos (\omega_c + \omega_s)t\right] \end{aligned} \tag{5}$$

The wave of Fig. 208 evidently contains three frequencies; the original carrier frequency and two others, one lying above and the other below the carrier frequency. These may be viewed as the lines of a frequency spectrum as shown in Fig. 209. As the modulating source will ordinarily consist of a band of audio frequencies, there will be any number of sum and difference

terms, depending upon the nature of ω_s. These sum and difference terms are called *side bands*. Consequently, if a group of audio frequencies extending up to 5000 cycles are used to modulate a carrier frequency of 10^6 cycles, a band of frequencies 10,000 cycles in width will be produced, extending from 995 to 1005 kc, with the carrier frequency located in the center of the band.

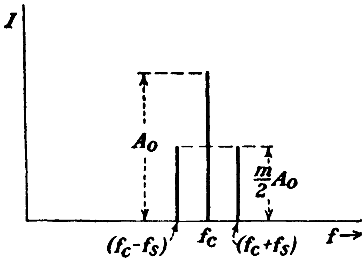


FIG. 209.—Carrier and side-band frequencies represented as the lines of a frequency spectrum.

The production of the three frequencies of Fig. 208 in the case of the alternator of the previous example, where a low-frequency alternating current was superimposed upon the normal direct-current excitation, may be understood from a consideration of Fig. 210. In (*a*), the armature is assumed to be rotating in a stationary magnetic field of constant strength, represented by the vector F_{dc}. The frequency of the e.m.f. induced in the armature will depend upon its speed and the number of poles. Thus, a two-pole machine rotating at 3600 r.p.m. would have a

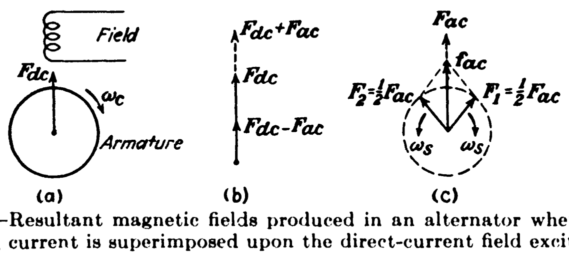


FIG. 210.—Resultant magnetic fields produced in an alternator when an alternating current is superimposed upon the direct-current field excitation.

"carrier" frequency f_c of 60 cycles induced in the armature. If an alternating "signal" frequency f_s of 5 cycles, for example, is then superimposed upon the direct-current excitation, the strength of the field will pulsate in magnitude, as shown in (*b*), assuming that the maximum value of the alternating-current field F_{ac} is less than the magnitude of the direct-current field F_{dc}. The pulsations caused by the alternating excitation may be thought of as having been produced by two oppositely rotating vectors, F_1 and F_2, each having a magnitude of $\frac{1}{2}F_{ac}$, whose resultant at any instant is f_{ac}, as in (*c*). The fields represented

by these two vectors glide around the air gap of the machine in opposite directions and with equal velocities of $\omega_s = 2\pi f_s$. With a clockwise rotation of the armature, the relative velocity with which the armature conductors cut the field F_1 will be $\omega_c - \omega_s$, while the field F_2 will be cut with a relative velocity of $\omega_c + \omega_s$. Three frequencies will consequently be present in the armature; the carrier f_c of 60 cycles, the lower side band produced by F_1 of 55 cycles, and the upper side band produced by F_2 of 65 cycles. If the direct-current field excitation is removed and only the 5-cycle alternating-current excitation remains, the 60-cycle carrier will disappear, leaving only the two side bands of 55 and 65 cycles. This simulates a "suppressed-carrier" system, which will be described in Sec. 135.

Frequency and phase modulation both produce side-band frequencies extending over a much wider range of frequencies than is the case with amplitude modulation.[1] A frequency-modulated wave contains a series of side bands which differ from the carrier frequency by integral multiples of the modulating frequency. If a carrier frequency of f_c cycles is frequency-modulated at a rate of f_s cycles per second, the side-band frequencies produced will be $f_c + f_s$, $f_c - f_s$, $f_c + 2f_s$, $f_c - 2f_s$, $f_c + 3f_s$, $f_c - 3f_s$, etc.

126. Energy Relations.—When the carrier wave is completely modulated ($m = 1$), the amplitude of the side bands is one-half that of the carrier. The energy is proportional to the square of the amplitude, and in a wave that is completely modulated the carrier represents $66\frac{2}{3}$ per cent of the total energy while the two side bands represent $33\frac{1}{3}$ per cent. The intelligence is conveyed entirely by the side bands so that it is important to employ as high a percentage of modulation as possible. The radio-telephone signal produced in a receiver by a strong carrier which is feebly modulated will be no louder than that of a weak carrier which is completely modulated. Furthermore, the strong carrier wave is capable of producing heterodyne interference over a much larger area than the completely modulated weaker signal, although the service areas of the two might be approximately equal.

The continuous power output with 100 per cent modulation is 1.5 times the output at zero modulation. The output at modula-

[1] H. RODER, Amplitude, Phase, and Frequency Modulation, *Proc. I.R.E.*, vol. 19, p. 2145, December, 1931; also see discussion of this paper in vol. 20, p. 884, May, 1932.

tions peaks in this case will be four times the unmodulated carrier output.

127. Methods of Obtaining Amplitude Modulation.—Amplitude modulation of a carrier wave can be secured in a variety of ways. Omitting telegraphic keying, the various methods may be classified as follows:

1. Absorption modulation.
2. Plate-circuit modulation.
3. Grid-circuit modulation.
4. Modulation by means of nonlinear impedances.

Most of these methods involve the use of a suitable vacuum tube, although the first and last can also be accomplished by other means. The method used will depend upon a variety of

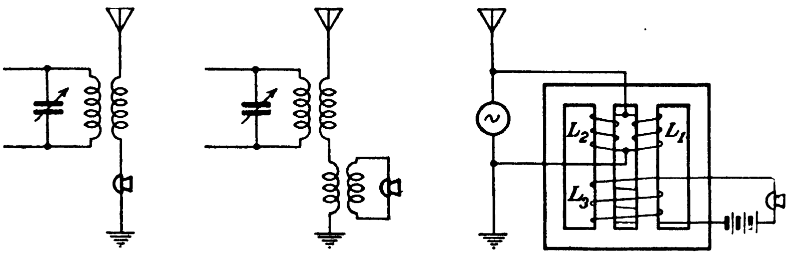


FIG. 211.—Circuits for absorption modulation.

factors, such as the amount of power to be handled, the percentage modulation desired, the cost of the necessary equipment, etc.

128. Absorption Modulation.—As its name implies, this method absorbs a portion of the high-frequency power by means of a voice-controlled impedance, such as a carbon-granule microphone. Typical circuits are shown in Fig. 211*a* and *b*. The limited current-carrying capacity of the microphone restricts the power that can be absorbed by the device to a maximum value of about 5 watts.

The magnetic modulator of Fig. 211*c* is capable of handling large amounts of power and was used to modulate the outputs of high-frequency alternators prior to the advent of vacuum tubes in this field.[2] Two similar coils L_1 and L_2 wound as shown are connected in parallel across the terminals of the alternator.

[2] E. F. ALEXANDERSON, A Magnetic Amplifier for Radio Telephony, *Proc. I.R.E.*, vol. 4, p. 101, April, 1916.

The voice coil L_3 is wound around both of the central legs of the iron core so that no high-frequency voltage will be induced in this circuit. The voice currents vary the flux density of the core and cause changes in the incremental permeability of the iron, so that the alternating-current impedance of coils L_1 and L_2 shunted across the alternator is made to vary. By proper magnetic design it is possible to produce a considerable change in the value of this impedance by a small change in the voice current. In this way the power absorbed by the device is made to vary in a manner dictated by the current through the microphone circuit. The moderate values of radio frequency produced by these alternators enable the magnetic properties of iron to be successfully employed.

129. Plate-modulated Oscillators.—This method, known as "constant current modulation" and also as "Heising modula-

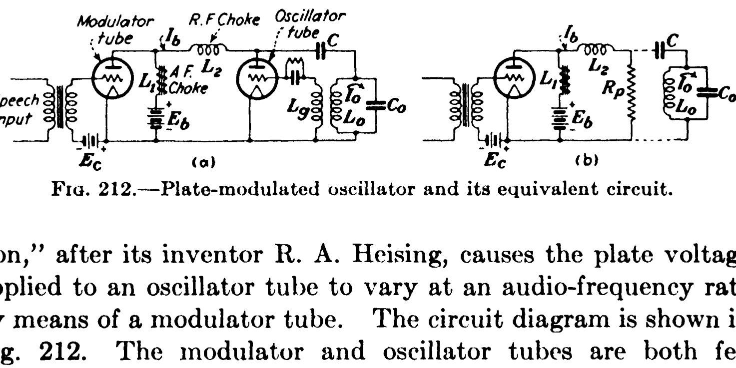


FIG. 212.—Plate-modulated oscillator and its equivalent circuit.

tion," after its inventor R. A. Heising, causes the plate voltage applied to an oscillator tube to vary at an audio-frequency rate by means of a modulator tube. The circuit diagram is shown in Fig. 212. The modulator and oscillator tubes are both fed from a common plate-supply voltage through the choke coil L_1. The inductive reactance of this coil should be large compared to the plate resistances of either tube at the lowest modulating frequencies to be used. The radio-frequency choke coil L_2 is to exclude radio-frequency currents from the modulator tube.

An inspection of the equivalent circuit reveals the modulator to be nothing but a choke-coupled, audio-frequency power amplifier, using shunt feed and having a useful load R_p. The tube operates as a Class A device and the theory developed in Chap. VII concerning power amplifiers is therefore directly applicable to modulators. Any power-amplifier tube capable of a large undistorted power output is satisfactory as a modulator.

The load resistance R_p of the oscillating tube is given by

$$R_p = \frac{E_b}{I_b} \tag{6}$$

and is the reciprocal of the slope of the E_b-I_b characteristic of the tube when oscillating.

The oscillator should be adjusted so that this characteristic is essentially linear, otherwise R_p will vary with the modulating voltage and the modulated output will be distorted as a result of this fluctuating load. A similar linear relation must also exist between the oscillatory tank current I_o and the plate voltage, or the modulated envelope of the carrier wave will be distorted. The voltage applied to the oscillator will be the plate-supply voltage E_b with the audio-frequency-voltage output of the modulator superimposed upon it. The grid excitation, grid bias, and the impedance of the tank circuit are adjusted until the desired linear relations are obtained. The adjustments can be checked by varying E_b from as small a value as will still permit the circuit to continue oscillating, to twice the normal value. The resultant values of I_o and I_b can then be plotted against E_b, as in Fig. 213, for the case of a modulated amplifier. The value of tank-circuit impedance required is usually somewhat greater than normal, resulting in a reduction of the plate efficiency. The value of E_b impressed upon the modulated oscillator is somewhat lower than the normal value used for unmodulated operation in order to avoid excessive plate heating on modulation peaks.

The average power output to the tank circuit when the wave is completely modulated is 1.5 times the unmodulated value. This increase in power is represented by the energy in the side bands, since the carrier power is not affected by the degree of modulation. The power associated with each side band is $m^2/4$ times the power associated with the carrier. This increase in power when modulated must be furnished by the alternating-current power output of the modulator tube. Thus, if the power represented by an unmodulated carrier wave is 100 watts, the average power will rise to 150 watts at 100 per cent modulation, 50 watts being furnished by the modulator tube and 100 watts from the direct-current supply to the oscillator. Since the side-band power varies as the square of the modulation factor, reducing the modulation to $m = 0.5$ would require only 12.5 watts from the modulator tube. The plate efficiency of the oscillator has been neglected in this discussion and consequently

the net input to the oscillator will be its output divided by its plate efficiency. To summarize these relations, the power required to generate the carrier wave is furnished by the direct-current plate supply, while the power required to produce the side bands is furnished from the alternating-current output of the modulator tube. The modulator output determined in this manner will be the *average* power output.

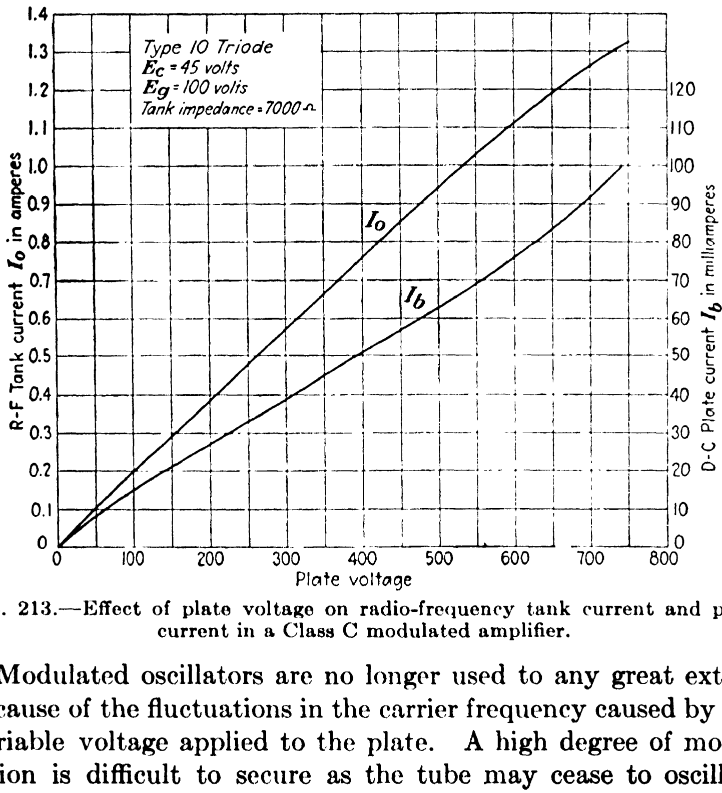


FIG. 213.—Effect of plate voltage on radio-frequency tank current and plate current in a Class C modulated amplifier.

Modulated oscillators are no longer used to any great extent because of the fluctuations in the carrier frequency caused by the variable voltage applied to the plate. A high degree of modulation is difficult to secure as the tube may cease to oscillate at low values of plate voltage.

130. Plate-modulated Amplifiers.—The disadvantages just mentioned are avoided if the modulation is applied to the plate circuit of a separately excited Class C amplifier instead of an oscillator. The equivalent circuit of Fig. 212*b* applies equally well to a modulated amplifier, since the only difference between an oscillator and an amplifier is in the source of grid excitation.

Linear relations between tank current I_o and plate current I_b with variable plate voltage are again necessary. These are

illustrated in Fig. 213 for a small triode. A higher value of tank-circuit impedance would have made the curve for I_o more nearly a straight line, but with some sacrifice in power output. The biasing voltage E_c should be approximately twice the value of cut-off. The output can be completely modulated by causing the plate voltage to vary from zero to twice the value of the direct-current plate-supply voltage. There is no danger in this case of stopping the oscillations because of low plate voltage, as is the case with an oscillator. It is impossible to swing the plate voltage applied to the amplifier between these limits with the modulator circuit of Fig. 212, since the modulator plate voltage would also have to vary between these same limits. The reason for this is apparent from a consideration of the I_p-E_p diagram of a Class A amplifier, such as Fig. 121. Modulator circuits which enable 100 per cent modulation to be secured will be described presently.

It will be noted in Fig. 212*b* that the audio-frequency output of the modulator is supplied to R_p through the radio-frequency choke coil L_2. The reactance of this choke coil should therefore be small compared to R_p at the highest modulating frequency to be used. This limits L_2 to a value no larger than that necessary to exclude radio-frequency currents from the modulator tube or the distributed capacitance of L_1. This is readily achieved when the carrier frequency used is high compared to the highest modulating frequencies. But at low carrier frequencies the choke required may offer appreciable impedance to the higher voice frequencies, with resultant frequency distortion. A parallel-resonant circuit tuned to the carrier frequency can be substituted for L_2 in such cases.

The blocking condenser C is effectively in parallel with R_p, as the tank-circuit impedance is negligible at audio frequencies. Consequently, the higher modulating frequencies tend to be by-passed around R_p if C is made too large. This item should also be made no larger than necessary.

These considerations as to the values of L_2 and C_2 apply equally well to modulated oscillators. The power amplifiers other than the modulated stage have no such limitations imposed upon the maximum size of the chokes and blocking condensers.

The power required to modulate an amplifier is determined in the same manner as described for modulated oscillators. The

audio-frequency power necessary is relatively large, particularly as m approaches unity. It is therefore advisable to modulate one of the earlier stages where the energy level is low. For example, if a 5000-watt carrier is to be completely modulated, the side-band power will be 2500 watts. Assuming the modulated amplifier to have a plate efficiency of 60 per cent, the required undistorted output of the modulator will be 4167 watts. An audio-frequency amplifier capable of furnishing this amount of power would be sizable and costly. When similar tubes are used for both modulator and amplifier, at least three or four tubes will be required to produce the power necessary to modulate a single amplifier tube, as the plate efficiency of the Class A modulator is rarely more than 20 per cent.

The Class B amplifier stages which follow the modulated stage should not employ too high a value of Q in their tank circuits, otherwise the tuning may be too sharp and the outer side-band frequencies may be discriminated against.

Either triodes or screen-grid tetrodes can be used in the modulated stage. When the former are used, some means of neutralization must be employed to prevent self-oscillation.

131. Modulator Coupling Circuits.—In order for an amplifier or an oscillator to be completely modulated, the plate voltage must vary from zero to twice the value of E_b, as illustrated in Fig. 186. As this extreme variation is not possible with the modulator tube acting as a Class A amplifier, the coupling arrangement of Fig. 214a can be used. A resistance R_1 is placed in series with the plate of the modulated amplifier, and of sufficient size to drop the direct-current plate voltage applied to a value just equal to the maximum amplitude of alternating-voltage output of the modulator. By operating the modulated amplifier at a lower plate voltage than the modulator tube it becomes possible to secure 100 per cent modulation. The voltage-dropping resistance R_1 is shunted by a large by-pass condenser C_1 which offers a path of low impedance for the audio-frequency output of the modulator. The reactance of this condenser should be small compared to R_p of the modulated amplifier for the lowest modulating frequencies to be used. This ordinarily requires a condenser from 5 to 20 μf, or larger.

Another method of securing complete modulation is to replace L_1 with a suitable autotransformer which steps up the audio-

frequency voltage output of the modulator to a value just equal to E_b on modulation peaks, as in Fig. 214b. The plate voltage applied to the modulated amplifier tube will then vary from

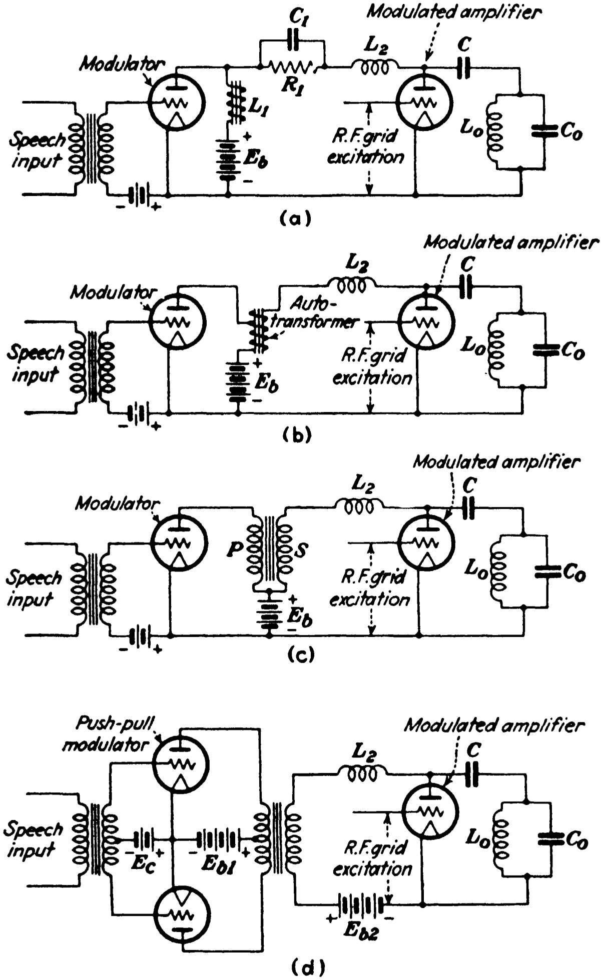


FIG. 214.—Various methods of coupling a modulator to the modulated amplifier to secure 100 per cent modulation.

zero to $2E_b$. This transformer also steps down the value of R_p as viewed from the modulator tube by the square of the ratio of transformation used. In this way it is possible to adjust the load impedance of the modulator tube to its optimum value by the

choice of the proper transformation ratio. In Fig. 212 the value of the load R_p is determined by the characteristics of the modulated amplifier. This value of R_p may be quite different from that required by the modulator for maximum undistorted output.

Instead of an autotransformer a two-winding transformer may be used, as in Fig. 214*c*. The direct current for the modulator flows through P, while that taken by the modulated amplifier flows through S. This method of coupling offers a decided advantage over the autotransformer in that the primary and secondary windings can be connected so that the direct-current ampere-turns oppose each other. This greatly reduces the direct-current flux density in the core and permits a much smaller cross section of iron to be used. In the previous methods of coupling the direct current flowing through L_1 was the sum of the two currents. This requires a relatively large core area in order to keep down the direct-current saturation and secure a reasonable value of incremental permeability from the core material. The determination of the inductance under these conditions has already been discussed in Sec. 29, Chap. III.

When a coupling transformer is used, it is no longer necessary to use a common source of plate-supply voltage for both modulator and amplifier. The two tubes, if desired, can now be operated from separate sources having the voltage values best suited to the types of tubes used. The modulator may use two tubes in push-pull, thereby obtaining the advantages of lower distortion and higher output that pertain to this mode of operation. The connections are shown in Fig. 214*d*. However, the direct-current ampere-turns on the two halves of the primary cancel each other, and hence are not available to oppose the ampere-turns of the secondary produced by the direct current taken by the modulated amplifier. The coupling transformer will accordingly require a larger core than the single-tube arrangement of Fig. 214*c*.

When two tubes are used in push-pull, they are very often operated Class B in order to obtain the relatively large output possible under this condition. The advantages and disadvantages of this method as compared to Class A operation have been already discussed in Chap. VII. The calculation of output and distortion in modulators is exactly the same as for power amplifiers, and it will not be repeated here. The problem of

direct-current saturation in the core of the push-pull coupling transformer is the same with either Class A or Class B operation.

Pentodes could also be used as modulators, but they are much more critical as to load impedance than triodes and their distortion is greater.

132. Grid-modulated Amplifiers.—Instead of varying the voltage applied to the plate of the modulated amplifier, it is possible to secure the same results by varying the magnitude of the C bias at an audio-frequency rate. The schematic circuit is shown in Fig. 215, together with the details of operation.

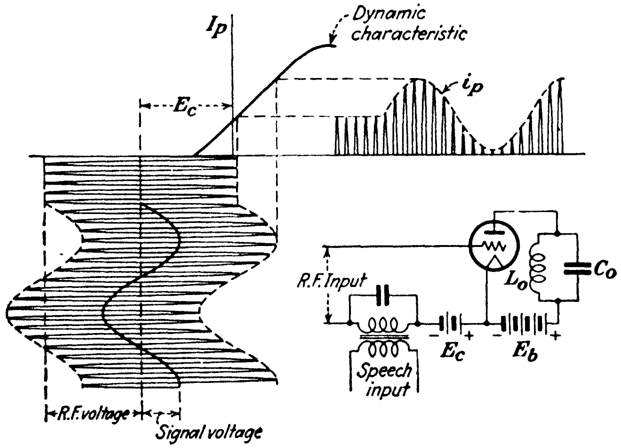


FIG. 215.—Schematic circuit and details of operation of a grid-modulated amplifier.

The tube is operated as a Class C amplifier with a biasing voltage E_c greater than cut-off. The tank-circuit impedance must be made large enough to insure a linear dynamic I_p-E_g characteristic over the operating range of grid voltages. The radio-frequency and signal voltages are connected in series and applied to the grid. The signal voltage cyclically adds to, and subtracts from, the fixed biasing voltage E_c, causing the amplitude of the plate-current impulses to rise and fall. The plate-current wave shapes will be similar to those of the Class B amplifier of Fig. 184, except that the angle $2\theta_1$ during which plate current flows will vary with the modulation. The mode of operation changes from an underexcited Class C amplifier when unmodulated, to a Class B amplifier on modulation peaks, assuming complete modulation.

The advantage of this method over plate modulation is that very little audio-frequency energy is required for complete modulation. The modulating source is only required to furnish a part of the grid-excitation losses of the modulated Class C amplifier in this case, instead of having to produce the power necessary to generate the side bands. This greatly reduces the size and cost of the audio-frequency-modulating equipment needed. However, the plate efficiency of the modulated amplifier is lower than with plate modulation. This is principally due to the lower average value of grid excitation applied. A linear relation between the tank current I_o and the modulating voltage is also somewhat more difficult to secure, although these adjust-

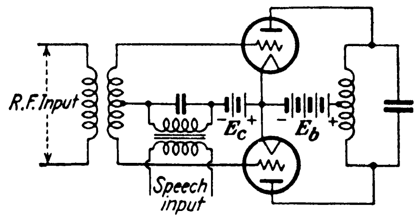


FIG. 216.—Schematic circuit of a push-pull grid-modulated amplifier.

ment difficulties are offset by the absence of a similar required relationship between E_b and I_b.

Grid modulation can be conveniently adapted to a push-pull connection as shown in Fig. 216. This method has been used with low-power broadcast transmitters.[3]

133. Modulation Due to a Nonlinear Impedance.—When a sine wave of voltage is impressed on a nonlinear impedance, such as a vacuum tube, new frequencies will be produced. The relationship between the current and voltage in any nonlinear device may be expressed in the form of a series as

$$i = i_o + ae + be^2 + ce^3 + \cdots \tag{7}$$

The constants in (7) may be evaluated by differentiation as discussed in Sec. 54, Chap. VI.

If a voltage of $e_g = E_g \sin \omega t$ is impressed on the grid of a tube having no C bias and zero external impedance in the plate circuit, the plate current will be given, from (24), Chap. VI, by

[3] A. W. KISHPAUGH, A Low-power Broadcast Transmitter, *Bell Lab. Rec.*, vol. 11, p. 37, October, 1932.

$$I_p = I_b + \frac{dI_p}{dE_g}E_g \sin \omega t + \frac{d^2I_p}{dE_g^2}\frac{E_g^2}{2}\sin^2 \omega t + \frac{d^3I_p}{dE_g^3}\frac{E_g^3}{6}\sin^3 \omega t + \cdots$$
$$= I_b + g_mE_g \sin \omega t + \frac{dg_m}{dE_g}\frac{E_g^2}{2}\left(\frac{1}{2} - \frac{1}{2}\cos 2\omega t\right) +$$
$$\frac{d^2g_m}{dE_g^2}\frac{E_g^3}{6}\left(\frac{3}{4}\sin \omega t - \frac{1}{4}\sin 3\omega t\right) + \cdots \quad (8)$$

From the standpoint of an amplifier the terms may be grouped as follows:

$$\left.\begin{array}{lll} I_p = I_b + \dfrac{dg_m}{dE_g}\dfrac{E_g^2}{4} + \cdots & & \text{Direct-current terms} \\ + \left(g_mE_g + \dfrac{d^2g_m}{dE_g^2}\dfrac{E_g^2}{8} + \cdots\right)\sin \omega t & & \text{Useful terms} \\ \left.\begin{array}{l} - \dfrac{dg_m}{dE_g}\dfrac{E_g^2}{4}\cos 2\omega t \\ - \dfrac{d^2g_m}{dE_g^2}\dfrac{E_g^2}{24}\sin 3\omega t \\ + \cdots \end{array}\right\} & & \text{Distortion terms} \end{array}\right\} \quad (9)$$

The direct-current terms following I_b account for the rise in the average value of the plate current from I_b to I_b'. With a linear characteristic, $dg_m/dE_g = 0$ and no rectification takes place. The distortion terms will likewise disappear.

If two frequencies, a carrier and a signal, are applied to the grid so that

$$E_g = E_1 \sin \omega_c t + E_2 \sin \omega_s t \quad (10)$$

the plate current in (8) becomes

$$I_p = I_b + g_m(E_1 \sin \omega_c t + E_2 \sin \omega_s t) +$$
$$\frac{g_m'}{2}(E_1^2 \sin^2 \omega_c t + 2E_1E_2 \sin \omega_c t \sin \omega_s t + E_2^2 \sin^2 \omega_s t)$$
$$+ \frac{g_m''}{6}(E_1^3 \sin^3 \omega_c t + 3E_1^2E_2 \sin^2 \omega_c t \sin \omega_s t + 3E_1E_2^2 \sin \omega_c t$$
$$\sin^2 \omega_s t + E_2^3 \sin^3 \omega_s t) + \cdots \quad (11)$$

where

$$g_m' = \frac{dg_m}{dE_g}, \qquad g_m'' = \frac{d^2g_m}{dE_g^2}$$

Tabulating the various terms as before, we get

$$
\left.\begin{aligned}
I_p &= I_b + \frac{g'_m}{4}(E_1^2 + E_2^2) && \text{Direct-current terms} \\
&+ \left(g_m E_2 + \frac{g''_m E_1^2 E_2}{4} + \frac{g''_m E_2^3}{8}\right)\sin \omega_s t && \text{Signal frequency} \\
&\left.\begin{aligned}&- \frac{g'_m E_2^2}{4}\cos 2\omega_s t \\ &- \frac{g''_m E_2^3}{24}\sin 3\omega_s t\end{aligned}\right\} && \text{Harmonics of signal frequency} \\
&\left.\begin{aligned}&- \frac{g''_m E_1 E_2^2}{8}\sin(\omega_c - 2\omega_s)t \\ &+ \frac{g'_m E_1 E_2}{2}\cos(\omega_c - \omega_s)t\end{aligned}\right\} && \text{Lower side band} \\
&+ \left(g_m E_1 + \frac{g''_m E_1 E_2^2}{4} + \frac{g''_m E_1^3}{8}\right)\sin \omega_c t && \text{Carrier} \\
&\left.\begin{aligned}&- \frac{g'_m E_1 E_2}{2}\cos(\omega_c + \omega_s)t \\ &- \frac{g''_m E_1 E_2^2}{8}\sin(\omega_c + 2\omega_s)t\end{aligned}\right\} && \text{Upper side band} \\
&\left.\begin{aligned}&+ \frac{g''_m E_1^2 E_2}{8}\sin(2\omega_c - \omega_s)t \\ &- \frac{g'_m E_1^2}{4}\cos 2\omega_c t \\ &- \frac{g''_m E_1^2 E_2}{8}\sin(2\omega_c + \omega_s)t\end{aligned}\right\} && \text{Second-order carrier terms} \\
&- \frac{g''_m E_1^3}{24}\sin 3\omega_c t && \text{Third harmonic of carrier}
\end{aligned}\right\} \quad (12)
$$

If more terms had been included in the series of (7) third-order side-band terms of the form $(3\omega_c \pm \omega_s)$ would have accompanied the third harmonic of the carrier. The upper and lower side bands of ω_c will include harmonics of the signal frequency. These harmonics will be absent if the characteristic of the tube follows a square law, *i.e.*, if the higher order terms above be^2 in (7) are negligible. In this case the carrier-frequency terms will be

$$
\begin{aligned}
i_c &= g_m E_1 \sin \omega_c t + \tfrac{1}{2} g'_m E_1 E_2 \cos(\omega_c - \omega_s)t \\
&\qquad - \tfrac{1}{2} g'_m E_1 E_2 \cos(\omega_c + \omega_s)t \\
&= g_m E_1\left(1 + \frac{g'_m}{g_m} E_2 \sin \omega_s t\right)\sin \omega_c t \\
&= g_m E_1 (1 + m \sin \omega_s t)\sin \omega_c t \qquad (13)
\end{aligned}
$$

where m is the modulation factor and is equal to

$$m = \frac{g'_m}{g_m} E_2 \tag{14}$$

Equation (13) is identical in form with (4), showing that amplitude modulation will occur whenever two or more different frequencies are impressed on a nonlinear impedance. The value of g'_m/g_m will be small in a vacuum tube so that E_2 must be relatively large to secure a value of m approaching unity.

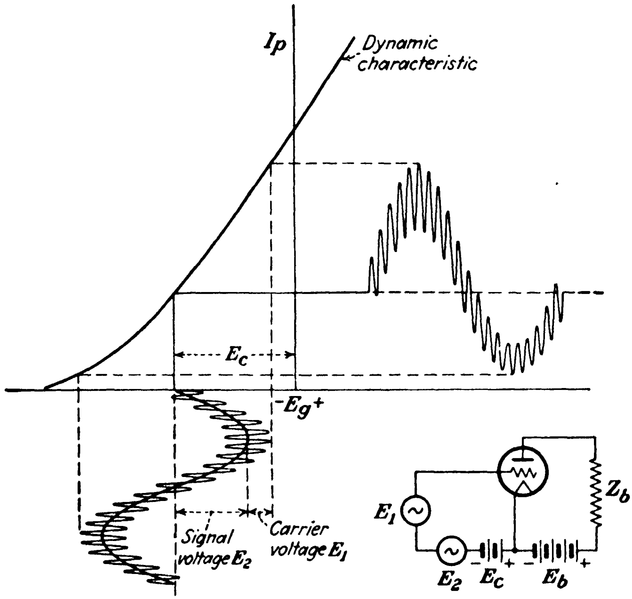


FIG. 217.—Schematic circuit and details of operation of a van der Bijl type modulator.

134. The van der Bijl Type of Modulated Amplifier.—A modulator using the principle of a nonlinear impedance, due to H. J. van der Bijl, is shown in Fig. 217. The carrier and signal-voltage sources are connected in series and applied to the grid of a negatively biased tube. As the operation of the device depends upon the curvature of the dynamic I_p-E_g characteristic, the load impedance Z_b in the plate circuit must not be too large because of the straightening effect it would have. It will be shown later that in the case where the load in the plate circuit is a pure resistance R_b, the modulated-power output will be a maximum when $R_b = \frac{1}{5}r_p$.

The output of the modulator contains numerous undesired frequencies as will be seen from (12). Suitable tuned circuits or a band-pass filter must therefore be incorporated in the output circuit which will only transmit the desired frequency terms. Modulators of this type have been used extensively in carrier-current telephony.[4]

135. Balanced Modulators for Carrier Suppression.—With a modulated wave the intelligence is transmitted only by the side bands and the carrier serves no useful purpose except at the receiving point. Here, the beats between the carrier and the two side bands produce the signal frequency. It would therefore be possible to suppress the carrier at the transmitter and transmit only the two side bands. At the receiving point the carrier frequency could be supplied by a local oscillator. However, if both side bands were transmitted, the missing carrier would have to be restored at the receiver at exactly the proper frequency and phase, which would be extremely difficult to do. But if only one side band is transmitted, the missing carrier can be supplied at the receiver to within perhaps 30 cycles of the correct value without producing objectionable distortion in the case of speech. With music the tolerance would be much lower.

By connecting two modulator tubes of the van der Bijl type in push-pull it is possible to suppress the carrier frequency and allow the two side bands to survive. One of the two side bands may then be removed from the output of the balanced modulator circuit by means of a suitable filter. In addition to other advantages, this method of single side-band transmission cuts in half the width of the frequency spectrum required for the transmission of intelligence. This is of considerable importance in carrier telephony which employs a frequency range extending from about 7 to 40 kc, as more communication channels are thereby provided.

The suppression of the carrier frequency may be accomplished in several ways. If an ordinary push-pull circuit is provided with two additional transformers at A and D, as in Fig. 218, there will

[4] For a comprehensive discussion of carrier current systems, see papers by E. H. Colpitts and O. B. Blackwell, Carrier Current Telephony and Telegraphy, *Trans. A.I.E.E.*, vol. 40, p. 205, 1921; and H. A. Affel, C. S. Demarest, and C. W. Green, Carrier Systems on Long Distance Telephone Lines, *Bell System Tech. Jour.*, vol. 7, p. 564, July, 1928.

be two places A and B where input voltages can be applied and two places C and D where the output power can be taken.

If the carrier and signal voltages are both applied to A, the two tubes are effectively in parallel with each other and no output voltage will appear across C. The output across D will contain all of the frequency terms listed in (12). If the signal E_2 is applied to B while the carrier E_1 is inserted at A, the side-band frequencies will appear across C, but the carrier will be suppressed as the two currents of this frequency are 180 degrees out of phase with each other in the two halves of the primary of transformer C. The currents in tube 1 due to E_2 will be given by (12). The currents in tube 2 will be given by the same equation, except that E_2 will now be negative, since E_2 causes the potential of one grid to rise

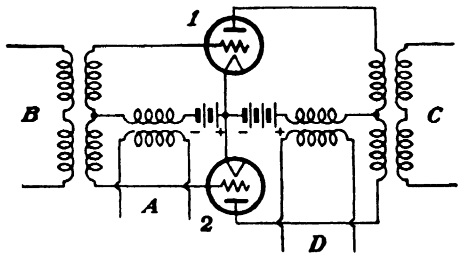


FIG. 218.—Balanced modulator circuit for carrier suppression.

while that of the other falls. If we reverse the sign of E_2 in (12) and designate the result as (12a), the frequency terms appearing across C will be (12) minus (12a). The frequency terms appearing across D will be (12) plus (12a). The carrier E_1 across A will cause the grid potentials of both tubes to rise and fall in unison so that the phase of E_1 is the same for both tubes and hence its sign will be the same in (12) and (12a).

Suppose the carrier and signal are interchanged so that E_1 is impressed across B and E_2 is inserted at A. The currents in tube 1 will be given by (12), as before. The currents in tube 2 will be given by the same expression, except that now the sign of E_1 must be reversed. Calling this latter equation (12b), the frequency terms appearing across C will now be (12) minus (12b), while those across D will be the sum of (12) and (12b).

The output-frequency terms obtained are summarized in Table I for the four possible input combinations, c and s designating carrier and signal, respectively. Combinations 3 and 4 are

usually the ones employed for carrier suppression. When 3 is used, the output transformer D is not needed, and in 4 both A and C can be omitted.

136. Grid-current Modulation.—Instead of making use of the nonlinear I_p-E_p characteristic as in the van der Bijl type of modulator, it is also possible to use the nonlinear relation between

TABLE I

	Signal in at	Carrier in at	Output at C	Output at D
1	A	A	0	$s, 2s, 3s, c, 2c, 3c, (c \pm s), (c \pm 2s), (2c \pm s)$
2	A	B	$c, 3c, (c \pm s), (c \pm 2s)$	$s, 2s, 3s, 2c, (2c \pm s)$
3	B	A	$s, 3s, (c \pm s), (2c \pm s)$	$c, 2c, 3c, 2s, (c \pm 2s)$
4	B	B	$c, 3c, s, 3s, (c \pm 2s) (2c \pm s)$	$2c, 2s, (c \pm s)$

the grid voltage and grid current. A considerable increase in both efficiency and side-band power output is obtained by this method.[5] Carrier and signal voltages are applied to the tube through a suitable impedance Z_c in series with the grid, as shown in Fig. 219. These voltages are large enough to cause grid current to flow. Owing to the curvature in the grid-current characteristic, modulation occurs in the grid circuit. These various modulation terms flow through Z_c and produce a voltage drop across it which is then amplified by the tube functioning in the ordinary manner as an amplifier. Filter circuits are provided in the output circuit to remove the undesired frequencies. While Fig. 219 shows only a single tube, grid-current modulators in practice are always used in one of the balanced circuit arrangements of Fig. 218 so as to suppress the

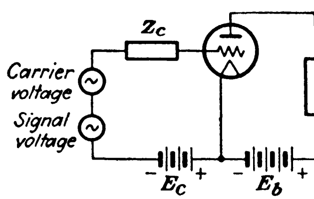


FIG. 219.—Schematic circuit of a grid-current modulator.

[5] E. PETERSON and C. R. KEITH, Grid Current Modulation, *Bell System Tech. Jour.*, vol. 17, p. 106, January, 1928

carrier frequency. They are chiefly used in carrier-current telephony.

137. Single-side-band Transmission.—As already mentioned, single-side-band transmission cuts the width of the frequency band required for the transmission of intelligence in half and thus provides more communication channels within a given range of frequencies. In radio communication another very important advantage offered is in the reduction of the radiated power required. With a completely modulated signal using the conventional method of transmission the carrier represents two-thirds of the total energy, while the two side bands represent the balance. For example, the energy in the carrier wave of a 1-kw. transmitter when unmodulated would be 1000 watts. When completely modulated, the carrier energy remains unchanged, but the total energy output rises to 1500 watts, the increase being due to the production of two 250-watt side bands. If the carrier and one side band had been suppressed at the source the same strength of received signal would have been produced by the radiation of a single 250-watt side band. Or, if all of the energy formerly employed was now devoted to the production of a single side band, a very much stronger signal would be received. The missing carrier would have to be replaced at the receiving station, but its energy need be no greater than that of the received carrier, had it been transmitted in the conventional manner.

Another advantage is in the reduction of selective fading when single-side-band transmission is used.[6]

Single-side-band transmission was invented by J. R. Carson[7] and is widely used in carrier-current telephony in order to conserve channel space. The reduction in the amount of energy which has to be transmitted also reduces the amount of cross talk between adjacent telephone lines.

The system is also used for the long-wave channels of the transatlantic radio telephone.[8] The elimination of the carrier and one side-band takes place at a low energy level and the output is then amplified through successive stages and impressed upon

[6] R. Bown, D. K. Martin and R. K. Potter, Some Studies in Radio Broadcast Transmission, *Proc. I.R.E.*, vol. 14, p. 57, February, 1926.

[7] U. S. patents No. 1,343,306, 1,343,307, and 1,449,382.

[8] R. A. Heising, Production of Single Side-band for Transatlantic Radio Telephony, *Proc. I.R.E.*, vol. 13, p. 291, June, 1925.

the antenna. The schematic arrangement of the various circuit elements is shown in Fig. 220. Double modulation is employed as it would be difficult and costly to construct an adjustable filter having a sharp enough cut-off frequency to transmit satisfactorily one side band and eliminate the other at frequencies in the vicinity of 60,000 cycles. An initial frequency of 33,700 cycles is accordingly introduced into a balanced modulator which eliminates the carrier. The modulator output passes through a band-pass filter which transmits the lower side band and eliminates the upper side band, which can readily be done at this low frequency. The output of this filter modulates a second carrier frequency in the vicinity of 89,200 cycles. This second carrier frequency is again suppressed in the second balanced modulator, the output of which contains the "signal" frequency of

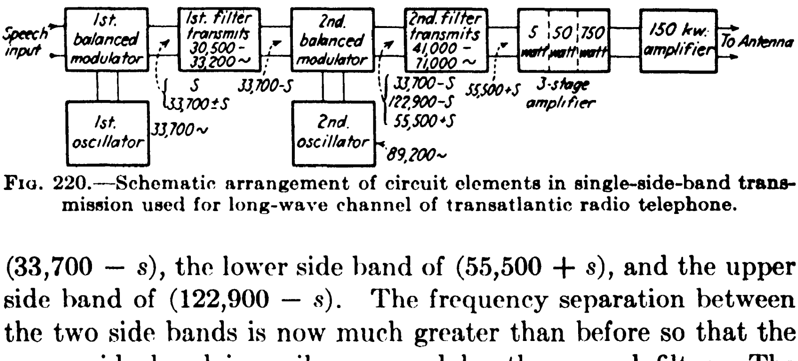


FIG. 220.—Schematic arrangement of circuit elements in single-side-band transmission used for long-wave channel of transatlantic radio telephone.

(33,700 − s), the lower side band of (55,500 + s), and the upper side band of (122,900 − s). The frequency separation between the two side bands is now much greater than before so that the upper side band is easily removed by the second filter. The transmitted band-width of this filter is purposely made wide so as to permit some adjustment to be made in the transmitted frequency. This change in frequency is brought about by adjusting the frequency of the second oscillator. The purpose is to be able to avoid the possibility of radio interference on the part of other long-wave transmitters.

Single-side-band transmission affords some measure of secrecy in the transmission as the speech is unintelligible unless the missing carrier is supplied. The production of a single side band also enables *inverted speech* to be produced by the simple expedient of supplying the missing carrier at the wrong edge of the side band. In this way the high speech frequencies become low and the low frequencies become high, making the result absolutely unintelligible. This has been suggested as a means of preserving

the secrecy of radio-telephone conversations. The selection of the upper side band by the first filter in Fig. 220 would have resulted in an inversion of the speech frequencies.

138. Other Types of Modulators.—Modulation can be produced in a tetrode by applying the carrier to the control grid and superimposing the modulating frequency upon the direct-current screen-grid potential, since the plate current is a function of the potentials of both grids. Serious amplitude distortion is produced in most tubes of this type if the degree of modulation is appreciable.

A more successful method is to insert the modulating voltage in the suppressor-grid circuit of a screen-grid type of pentode, as shown in Fig. 221. This grid is biased negatively by a moderate amount and swings positive on modulation peaks, during which time suppressor-grid current flows. The power represented by this flow of current has to be furnished by the modulating source, but it is very much smaller in amount than with plate-modulated amplifiers. No current flows when the suppressor-grid becomes slightly negative and consequently the modulating source must be designed to have good voltage regulation if amplitude distortion is to be avoided. The conditions are similar to the excitation demands of a triode amplifier when operated with insufficient negative grid bias. Owing to the much higher amplification of pentodes the radio-frequency excitation voltage that has to be applied to the control grid is much smaller than with similar triodes. The latter require power for grid-excitation purposes which is usually about 10 per cent of the plate-input power. The lower excitation voltage required by pentodes reduces the required excitation power to about 2 or 3 per cent of the plate input.

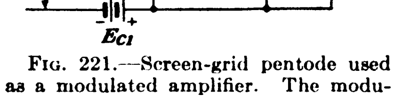


FIG. 221.—Screen-grid pentode used as a modulated amplifier. The modulating frequency is introduced into the suppressor-grid circuit.

Complete modulation can be obtained, although some amplitude distortion is apt to be produced at high levels of modulation. Other types of pentodes can also be used, but the shielding effect of the screen grid in those designed as audio-frequency power

amplifiers is usually insufficient to prevent self-oscillation at radio frequencies without the use of some form of neutralizing circuit.

139. Plate Current Expressed as an Infinite Series.—Equation (12) was derived for a triode having no external impedance in the plate circuit. The frequency terms present are unaffected by the plate load, but the amplitudes will obviously be altered. An expression for the plate current has already been derived in Sec. 56, Chap. VI, but it is valid only for the straight portion of the characteristic curve of the tube, or for values of applied grid voltage small enough to consider the working portion of the characteristic to be linear. This assumption can no longer be made in several of the types of modulators just considered, as their mode of operation definitely depends upon the curvature of the characteristic. A number of detectors or "demodulators" also depend upon this same feature.[9]

The plate current of a triode is some function of both the grid and plate voltages and may be expressed mathematically by

$$I_p = f(E_p, E_g) \tag{15}$$

where $I_p = I_b + i_p$

$$E_g = E_c + e_g$$
$$E_p = E_b + e_p \tag{16}$$

This expression can be expanded by the use of Taylor's series for a function of two independent variables which is of the form[10]

$$f(x, y) = f(a, b) + \left\{(x - a)\frac{\partial}{\partial x} + (y - b)\frac{\partial}{\partial y}\right\}_0 f + \frac{1}{\underline{|2}}\left\{(x - a)\frac{\partial}{\partial x} + (y - b)\frac{\partial}{\partial x}\right\}_0^2 f + \cdots + \frac{1}{\underline{|n}}\left\{(x - a)\frac{\partial}{\partial x} + (y - b)\frac{\partial}{\partial y}\right\}_0^n f + R \tag{17}$$

where x and y are the independent variables and the expansion is about the point $x = a$, $y = b$, where a and b are arbitrary constants. Expanding (17), we have

[9] J. R. CARSON, A Theoretical Study of the Three-element Vacuum Tube, *Proc. I.R.E.*, vol. 7, p. 187, April, 1919.

[10] F. S. WOODS, "Advanced Calculus," p. 86, Ginn and Company, Boston.

$$f(x, y) = f(a, b) + (x - a)\frac{\partial f}{\partial x} + (y - b)\frac{\partial f}{\partial y} + \frac{1}{\underline{|2}}\left\{(x - a)^2\frac{\partial^2 f}{\partial x^2} + 2(x - a)(y - b)\frac{\partial^2 f}{\partial x \partial y} + (y - b)\frac{\partial^2 f}{\partial y^2}\right\}$$
$$+ \frac{1}{\underline{|3}}\left\{(x - a)^3\frac{\partial^3 f}{\partial x^3} + 3(x - a)^2(y - b)\frac{\partial^3 f}{\partial x^2 \partial y} + 3(x - a)(y - b)^2\frac{\partial^3 f}{\partial x \partial y^2} + (y - b)^3\frac{\partial^3 f}{\partial y^3}\right\} + \cdots \quad (18)$$

In this expression a is any fixed value of x and b is any fixed value of y. The derivatives are all to be evaluated at the point where $x = a$ and $y = b$.

Letting

$$\begin{aligned} x &= E_p \\ y &= E_g \\ a &= E_b \\ b &= E_c \end{aligned}$$

the expansion of (15) becomes

$$I_p = f(E_p, E_g) = f(E_b + e_p, E_c + e_g) = I_b + i_p$$

$$I_b + i_p = f(E_b, E_c) + e_p\frac{\partial I_p}{\partial E_p} + e_g\frac{\partial I_p}{\partial E_g} + \frac{1}{2}\left\{e_p^2\frac{\partial^2 I_p}{\partial E_p^2} + 2e_pe_g\frac{\partial^2 I_p}{\partial E_p \partial E_g} + e_g^2\frac{\partial^2 I_p}{\partial E_g^2}\right\} + \frac{1}{6}\left\{e_p^3\frac{\partial^3 I_p}{\partial E_p^3} + 3e_p^2e_g\frac{\partial^3 I_p}{\partial^2 E_p \partial E_g} + 3e_pe_g^2\frac{\partial^3 I_p}{\partial E_p \partial E_g^2} + e_g^3\frac{\partial^3 I_p}{\partial E_g^3}\right\} + \cdots \quad (19)$$

Since

$$\left.\begin{aligned} I_b &= f(E_b, E_c) \\ g_m &= \frac{\mu}{r_p} \\ \frac{\partial I_p}{\partial E_g} &= \mu\frac{\partial I_p}{\partial E_p} \end{aligned}\right\} \quad (20)$$

the alternating component i_p of the plate current, if μ can be regarded as a constant,[11] becomes

[11] For cases where the variation in μ must be considered, see papers by F. B. Llewellyn, Operation of Thermionic Vacuum Tube Circuits, *Bell System Tech. Jour.*, vol. 5, p. 433, July, 1926; E. Peterson and H. P. Evans, Modulation in Vacuum Tubes Used as Amplifiers, *Bell System Tech. Jour.*, vol. 6, p. 442, July, 1927.

$$i_p = e_p\frac{\partial I_p}{\partial E_p} + \mu e_g\frac{\partial I_p}{\partial E_p} + \frac{1}{2}\left\{e_p^2\frac{\partial^2 I_p}{\partial E_p^2} + 2e_pe_g\frac{\partial}{\partial E_p}\left(\mu\frac{\partial I_p}{\partial E_p}\right) + \mu^2e_g^2\frac{\partial^2 I_p}{\partial E_p^2}\right\}$$
$$+ \frac{1}{6}\left\{e_p^3\frac{\partial^3 I_p}{\partial E_p^3} + 3e_p^2e_g\frac{\partial}{\partial E_p^2}\left(\mu\frac{\partial I_p}{\partial E_p}\right) + 3e_pe_g^2\frac{\partial}{\partial E_p}\left[\mu\frac{\partial}{\partial E_p}\left(\mu\frac{\partial I_p}{\partial E_p}\right)\right] + \mu^3e_g^3\frac{\partial^3 I_p}{\partial E_p^3}\right\} + \cdots$$
$$= (e_p + \mu e_g)\frac{\partial I_p}{\partial E_p} + \frac{1}{2}(e_p + \mu e_g)^2\frac{\partial^2 I_p}{\partial E_p^2} + \frac{1}{6}(e_p + \mu e_g)^3\frac{\partial^3 I_p}{\partial E_p^3} + \cdots \quad (21)$$

Evaluating the derivatives, we get

$$\left.\begin{aligned}
\frac{\partial I_p}{\partial E_p} &= g_p = \frac{1}{r_p} \\
\frac{\partial^2 I_p}{\partial E_p^2} &= \frac{\partial\frac{1}{r_p}}{\partial E_p} = -\frac{1}{r_p^2}\frac{\partial r_p}{\partial E_p} \\
\frac{\partial^3 I_p}{\partial E_p^3} &= \frac{\partial}{\partial E_p}\left(-\frac{1}{r_p^2}\frac{\partial r_p}{\partial E_p}\right)^* = \frac{2}{r_p^3}\left(\frac{\partial r_p}{\partial E_p}\right) - \frac{1}{r_p^2}\frac{\partial^2 r_p}{\partial E_p^2}
\end{aligned}\right\} \quad (22)$$

Substituting the values of (22) in (21) gives

$$i_p = \frac{\mu e_g + e_p}{r_p} - \frac{(\mu e_g + e_p)^2}{2r_p^2}\frac{\partial r_p}{\partial E_p} + \frac{(\mu e_g + e_p)^3}{6}\left\{\frac{2}{r_p^3}\left(\frac{\partial r_p}{\partial E_p}\right)^2 - \frac{1}{r_p^2}\frac{\partial^2 r_p}{\partial E_p^2}\right\} + \cdots \quad (23)$$

If the plate circuit contains a resistance R_b, e_p is given by

$$e_p = -i_pR_b$$

and (23) becomes

$$i_p = \frac{\mu e_g - i_pR_b}{r_p} - \frac{(\mu e_g - i_pR_b)^2}{2r_p^2}r_p' + \frac{(\mu e_g - i_pR_b)^3}{6}\left[\frac{2}{r_p^3}r_p'^2 - \frac{r_p''}{r_p^2}\right] + \cdots \quad (24)$$

where

$$r_p' = \frac{\partial r_p}{\partial E_p} \qquad \text{and} \qquad r_p'' = \frac{\partial^2 r_p}{\partial E_p^2}$$

Since i_p can be represented by the power series

* $d(uv) = udv + vdu$.

$$i_p = c_1e_g + c_2e_g^2 + c_3e_g^3 + \cdots \tag{25}$$

an explicit expression for i_p can be obtained by inserting (25) for i_p in (24) and equating like powers of e_g, term by term, and solving for the various values of c. Enough terms must be used in (25) so that the coefficients of the powers of e_g will be correct for the highest power of i_p that we are concerned with. From (25)

$$\left.\begin{aligned} i_p^2 &= c_1^2e_g^2 + 2c_1c_2e_g^3 + (2c_1c_3 + c_2^2)e_g^4 + \cdots \\ i_p^3 &= c_1^3e_g^3 + 3c_1^2c_2e_g^4 + \cdots \end{aligned}\right\} \tag{26}$$

Expanding (24), and inserting (25) and (26) and equating like powers of e_g, we get

$$c_1e_g = \frac{\mu e_g}{r_p} - \frac{R_b}{r_p}c_1e_g$$

from which

$$c_1 = \frac{\mu}{r_p + R_b} \tag{27}$$

Equating like powers of e_g^2, we get

$$c_2e_g^2 = -\frac{R_b}{r_p}c_2e_g^2 - \frac{\mu^2e_g^2r'_p}{2r_p^2} + \frac{\mu e_g^2R_br'_p}{r_p^2}c_1 - \frac{R_b^2r'_p}{2r_p^2}c_1^2e_g^2$$

from which

$$c_2 = -\frac{\mu^2r_pr'_p}{2(r_p + R_b)^3} \tag{28}$$

In a similar manner c_3 is found to be

$$c_3 = \frac{\mu^3r_p}{6(r_p + R_b)^5}[(2r_p - R_b)r_p'^2 - r_pr''_p(r_p + R_b)] \tag{29}$$

Substituting the values of these constants in (25), the expression for the alternating component of the plate current of a triode having a constant value of μ and with a resistance load in its plate circuit is

$$i_p = \frac{\mu e_g}{r_p + R_b} - \frac{\mu^2e_g^2r_pr'_p}{2(r_p + R_b)^3} + \frac{\mu^3e_g^3r_p}{6(r_p + R_b)^5}[(2r_p - R_b)r_p'^2 - r_pr''_p(r_p + R_b)] + \cdots \tag{30}$$

The first term in (30) will be recognized as the useful or fundamental component of plate current in the case of an amplifier, while the other terms represent distortion. The effect of large values of R_b upon the reduction in size of these distortion terms

is quite evident. The second term is responsible for the production of second harmonic and modulation, the latter resulting in the production of the sum and difference frequencies when two or more frequencies are impressed on the grid of an amplifier. This is often called "cross modulation" and is responsible for the production of cross talk in the multichannel amplifiers of carrier telephony. In addition to the modulation caused by the curvature of the characteristic, the variation in μ caused by the variations in grid and plate voltages is also responsible for an appreciable portion of the undesired modulation produced, particularly when R_b is made large compared with r_p. When this is taken into account, the second term of (30) has been shown by Llewellyn[11] to be

$$-\frac{1}{2}\left[\frac{\mu^2 r_p r_p'}{(r_p + R_b)^3} - \frac{2r_p \dfrac{\partial \mu}{\partial E_g}}{(r_p + R_b)^2}\right] e_g^2 \tag{31}$$

Cross modulation may occur in radio-frequency amplifiers and cause stations to be heard at settings of the tuning dial which have no apparent relation to the assigned frequency of the stations thus heard. This is usually the result of insufficient selectivity on the part of the tuned circuits preceding the first amplifying tube. For example, suppose one broadcasting station is transmitting on a frequency of 1300 kc, and another on 700 kc. If these two frequencies both manage to reach the grid of the first amplifying tube, sum and difference frequencies of 2000 kc and 600 kc, among others, will be present in its output if modulation occurs due to a nonlinear characteristic. If the receiving set is then tuned to 600 kc (or 2000 kc, if the tuning range permits), both stations will be heard, as the subsequent amplifying tubes will amplify this frequency. If the field strength of one station is much greater than the other, owing to greater power or closer proximity, the signals from the stronger station will mask those of the other. Any number of similar combinations can occur, which would result in serious interference. Consequently, the possibility of modulation occurring in the first tube must be guarded against in the design of the amplifier.

The second term of (30) becomes the useful one in the case of those types of modulators and detectors which depend for their

[11] *Loc. cit.*

operation upon the curvature of the I_p-E_g characteristic. As mentioned previously the load resistance R_b in the plate circuit will result in reduced curvature of the characteristic. The magnitude of the modulated-voltage output will be given by

$$e_m = -\frac{\mu^2 e_g^2 r_p r_p'}{2(r_p + R_b)^3} R_b \tag{32}$$

By differentiating this expression with respect to R_b, the modulated voltage is found to be a maximum when

$$R_b = \tfrac{1}{2} r_p \tag{33}$$

The modulated power output is equal to

$$P_m = \left[\frac{\mu^2 e_g^2 r_p r_p'}{2(r_p + R_b)^3}\right]^2 R_b \tag{34}$$

which will be a maximum when

$$R_b = \tfrac{1}{5} r_p \tag{35}$$

The choice of load resistance will therefore depend upon whether maximum voltage or maximum power output is sought.

Problems

1. A type 10 triode is used as a modulated amplifier and has the characteristics shown in Fig. 213. This tube is to be modulated by a type 50 tube whose characteristics are given in the following table. The modulating

CHARACTERISTICS OF TYPE 50 TRIODE

$E_c = 0$		$E_c = -70$		$E_c = -140$	
E_p	I_p*	E_p	I_p*	E_p	I_p*
50	11.5	300	12.0	500	3.3
100	31.0	325	19.5	525	6.6
150	55.5	350	28.0	550	12.0
200	85.0	375	38.3	575	18.4
250	119.0	400	50.0	600	26.0
		425	62.5	625	35.0

* Milliamperes.

circuit used is the same as Fig. 212 with $E_b = 350$ volts. Neglecting the resistance of the choke L_1 and assuming its inductance to be very large, what is the maximum percentage of modulation that can be obtained for a maximum value of signal of 70 volts impressed on the grid of the modulator tube? The C bias impressed on the modulator tube is -70 volts.

2. What is the power output of the modulator in Problem 1? What is the peak, average, and unmodulated power output to the tank circuit of the modulated amplifier, assuming a constant plate efficiency of 60 per cent?

3. If two type 50 tubes had been connected in parallel in Problem 1, what percentage of modulation would have been obtained?

4. If the circuit of Fig. 214*a* had been used in Problem 1, what value of R_1 would have been required to secure 100 per cent modulation? What would be the average value of plate voltage supplied to the modulated amplifier under this condition?

5. A transformer-coupled circuit similar to Fig. 214*c*, but with separate sources of plate-supply voltage, is used to modulate the tube of Problem 1. The plate potential of the modulator tube is raised to 400 volts, the C bias remaining at -70 volts. The plate-supply potential of the modulated amplifier remains at 350 volts. If the primary of the coupling transformer has 2000 turns, how many secondary turns are required to offer a load impedance of 4000 ohms to the modulator tube? What will be the power output of the modulator tube at fundamental frequency if the maximum signal voltage applied to its grid is 70 volts? What percentage modulation will be obtained with this signal voltage?

6. In Problem 5 what will be the direct-current ampere-turns acting on the core if the primary and secondary connections are such as to make these a minimum? How does this value compare with the autotransformer of Fig. 214*b*, assuming the primary turns and transformation ratio to be the same in each case?

7. The circuit of Fig. 214*d* is used to modulate the tube of Problem 1. Two type 50 tubes are used in push-pull with a plate supply of 400 volts and a negative C bias of -70 volts. The output transformer has 3000 primary turns with a center tap at 1500 turns. How many secondary turns are required to produce a plate-to-plate load of 5000 ohms, neglecting transformer losses? What will be the power output of the modulator with a grid-to-grid voltage of 140 volts, maximum? Can complete modulation be obtained with this arrangement?

8. A type 101-D telephone repeater tube is to be used as a van der Bijl type of modulator, as in Fig. 217. The tube constants are: $\mu = 5.84$, $r_p = 6400$ ohms, $r'_p = -61.3$. The load resistance R_b is adjusted for maximum side-band-voltage output. The carrier and signal frequencies impressed on the grid are $E_1 = 1 \sin 200{,}000t$ and $E_2 = 8 \sin 5000t$, respectively. Determine the magnitude of the carrier and side-band voltages across R_b, neglecting higher order terms above e_g^2. What is the percentage of modulation obtained?

9. Repeat Problem 8 using the same value of R_b as before but assume that μ varies with the impressed signal so that $\partial\mu/\partial E_g = 0.05$. What is the percentage of modulation now obtained?

10. The equation of a modulated wave is given by

$$i = I(1 + m \sin \omega_s t) \sin \omega_c t$$

Derive an expression for the effective value of the current in terms of I and m.

CHAPTER XII

VACUUM-TUBE DETECTORS

140. Operation of Detectors.—Detection or demodulation is the process whereby the signal is extracted from the impressed modulated wave. This is accomplished by the use of some nonlinear device which rectifies the wave and causes the average value of the output current to vary in accordance with the modulated envelope of the impressed wave. Formerly, the rectifying properties of certain mineral crystals such as galena, zincite, silicon, and others were extensively used as detectors, but these have since been practically all replaced by the thermionic vacuum tube which is more reliable and does not require critical adjustments. The first vacuum-tube detector was the two-element "Fleming valve." This was followed a few years later by the "audion" of de Forest, who introduced a third electrode in the form of a grid which greatly increased the sensitivity of the device. It is interesting to note that the diode is again being extensively used as a detector in modern broadcast-receiving sets. High sensitivity on the part of the detector is no longer of vital importance, as any deficiency in this item is readily taken care of by increased gain in the radio-frequency amplifier which precedes it. Consequently, detectors of comparatively low sensitivity and which are relatively free from distortion are more satisfactory in receiving sets designed for entertainment purposes than more sensitive detectors having somewhat greater distortion.

Rectification in the case of triodes may be accomplished by making use of the nonlinear portions of either the I_p-E_g or the I_g-E_g characteristics. When the amplitude of the impressed wave is small, the working portion of the characteristic may be adequately expressed by the first two terms of a power series

$$i = c_1e + c_2e^2 + \cdots \tag{1}$$

as discussed in Chap. XI. But as the impressed voltage becomes

larger, the higher order terms in the series become increasingly important and the computations involved, as additional terms are included, soon become laborious beyond the point of usefulness for ordinary purposes of investigation. Detector performance at high signal voltages can therefore usually be determined more readily by experimental methods. With small signal voltages the power series of (1) yields results sufficiently accurate for most purposes.

The rectified output with small signal voltages is proportional to the square of the impressed voltage and detectors operating in this manner are termed *square-law detectors*. They have the disadvantage of producing distortion which increases as the square of the percentage modulation of the received signal, and hence are unsatisfactory for the reception of broadcast entertainment if m is much greater than 0.5.

If the signal voltage applied to the detector is fairly large, it is possible to bias the tube to approximately cut-off and cause the circuit to operate in much the same manner as a Class B amplifier. The rectified output in this case is almost directly proportional to the impressed signal voltage and the distortion is greatly reduced. Detectors operating in this manner are called *linear detectors* or *power detectors*. Their sensitivity is lower than the usual type of square-law detector, but this is more than offset by the reduced distortion.

141. Plate Detection.—Rectification may be accomplished by making use of the curvature in the I_p-E_g characteristic of the tube. This is shown graphically in Fig. 222 for a modulated signal applied to the grid of a negatively biased triode which has a pure resistance in its plate circuit. The variation in the average value of plate current is shown by the dotted curve i_s. It will be seen that the variations in i_s follow the modulated envelope of the impressed wave.

The value of i_s for a pure resistance load R_b in the plate circuit is given by the second term of (30) in Chap. XI and is

$$i_s = \frac{-\mu^2 r_p r'_p}{2(r_p + R_b)^3} e_g^2 \tag{2}$$

assuming the impressed voltage e_g to be small so that the higher order terms in (30) may be neglected. The negative sign of (2) may be ignored when only the magnitude of the rectified-signal

current is of interest. The factor multiplying e_g^2 is called the *detection coefficient*.

If a modulated wave is impressed of the form

$$e_g = E(1 + m \sin \omega_s t) \sin \omega_c t \tag{3}$$

the plate current due to the second term of (30) becomes

$$i_p = \frac{-\mu^2 r_p r_p' E^2}{2(r_p + R_b)^3}(1 + m \sin \omega_s t)^2\left(\frac{1}{2} - \frac{1}{2}\cos 2\omega_c t\right) \tag{4}$$

where E is the maximum amplitude of the unmodulated carrier voltage applied to the grid.

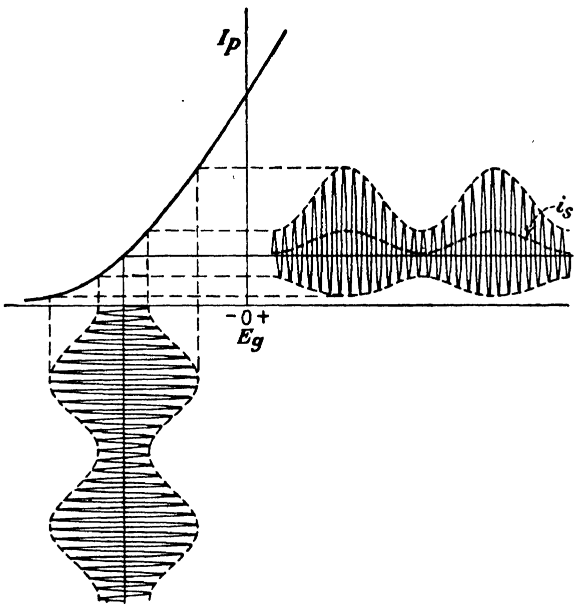


FIG. 222.—Plate detection in a triode resulting from curvature of the I_p-E_g characteristic.

The terms of audio frequency in (4) are

$$i_s = \frac{-\mu^2 r_p r_p' E^2}{2(r_p + R_b)^3}\left(m \sin \omega_s t - \frac{m^2}{4}\cos 2\omega_s t\right) \tag{5}$$

It will be observed from (5) that the output contains a second harmonic term having an amplitude which is proportional to the square of the modulating factor m. Consequently, a square-law detector will introduce 25 per cent distortion in the form of second harmonic if the received signal is completely modulated.

This is of little consequence in receivers designed for the reception of telegraphic code signals, but it would be excessively large in receiving sets designed for entertainment purposes. In addition, terms involving the sums and differences of the various audio frequencies will also be present, as pointed out in (4), Chap. VII. These sum and difference terms do not appear in (5) as only a single modulating frequency of ω_s has been assumed for simplicity, whereas actually there would be an entire band of frequencies. The detectors used in the older types of broadcast receivers were practically all of the square-law type, although they did not usually employ plate detection. At the time they were designed the average percentage of modulation used at broadcast stations was low, scarcely ever exceeding 50 per cent on modulation peaks, and the distortion produced by the detector was not very serious. Modern broadcast transmitters are practically all capable of being completely modulated so as to obtain the maximum possible service area, and, as a result, they will produce noticeable distortion in receiving sets equipped with square-law detectors.

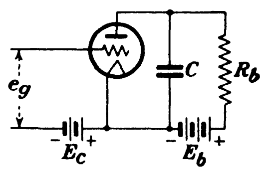


FIG. 223.—Circuit of a plate-curvature detector.

Equation (5) assumes a pure resistance load in the plate circuit of the detector, whereas in practice the load is always by-passed by a condenser in order to provide a path of low impedance for the radio-frequency components of plate current, as shown in Fig. 223. This increases the sensitivity of the device as a detector by reducing the magnitude of the denominator of the detection coefficient in (2). The impedance in the plate circuit will no longer be a pure resistance R_b, which was the condition assumed in the derivation of (30). It will not be correct to substitute the complex expression for the plate impedance in place of R_b in the preceding equations, since both the magnitude and the phase angle of the impedance will be functions of the frequency.

An expression for the plate current when the plate circuit contains an impedance can be obtained from (24) and (25) of Sec. 139, as before, by substituting the complex expression for the impedance in place of R_b in (24). An explicit expression for i_p can then be obtained by inserting (25) for i_p in (24) and equating like powers of e_g *which are of like frequency*, term by term, and

solving for the various values of c.[1] The complete expression for the plate current is somewhat involved when the applied grid voltage contains more than one frequency, as a different value of plate-circuit impedance is offered to each frequency.

Assuming a modulated signal having the equation given by (3), it can be shown that the terms of radian frequency ω_s in the plate circuit of Fig. 223 will be

$$i_s = \frac{-\mu^2 r_p r_p'}{2}\left[\frac{\frac{1}{2}mE^2}{(r_p + Z_c)(r_p + Z_{c-s})(r_p + Z_s)} + \frac{\frac{1}{2}mE^2}{(r_p + Z_c)(r_p + Z_{c+s})(r_p + Z_s)}\right] \tag{6}$$

The subscripts of the impedances in the denominator of the bracketed terms indicate that the impedance in the plate circuit is to be evaluated at that particular frequency.

A second harmonic of the signal frequency will also be produced, which will be

$$i_{2s} = \frac{\mu^2 r_p r_p'}{2(r_p + Z_{c+s})(r_p + Z_{c-s})(r_p + Z_{2s})} \frac{m^2}{4} E^2 \tag{7}$$

Additional distortion terms made up of the sums and differences of the various modulating frequencies will also be produced when other modulating frequencies are present in addition to ω_s.

When the plate of the detector is by-passed by the condenser C, as is practically always the case, the impedance offered by the external plate circuit to currents of carrier frequency will be substantially zero, so that $Z_c = 0$, $Z_{c+s} = 0$, $Z_{c-s} = 0$, and (6) becomes

$$i_s = \frac{-\mu^2 r_p' m E^2}{2r_p(r_p + Z_s)} \tag{8}$$

where Z_s is the vector expression for the impedance of C and R_b in parallel at the radian frequency ω_s. If the reactance of C is very large at the audio frequency in question, Z_s may be replaced by R_b.

An audio-frequency transformer can be substituted in place of R_b in Fig. 223, in which case Z_s will then be the impedance of the transformer primary in parallel with C. The value of C

[1] See McILWAIN and BRAINERD, "High Frequency Alternating Currents," pp. 125*ff*., John Wiley & Sons, Inc., 1931; or paper by F. B. LLEWELLYN, *Bell System Tech. Jour.*, vol. 5, p. 433, July, 1926.

should be large enough to by-pass the radio-frequency components of the plate current and at the same time offer a very high impedance to the audio-frequency terms. The by-passing effect of this condenser increases for the higher values of audio frequency so that some frequency distortion will be produced, causing the output voltage of the detector, which is

$$E_d = i_s Z_s \tag{9}$$

to be reduced for the higher audio frequencies. Greater difficulty is experienced in separating the radio- and audio-frequency components as the carrier frequency is lowered.

Plate detectors have not been very widely used in radio communication as they are not so sensitive as those using a

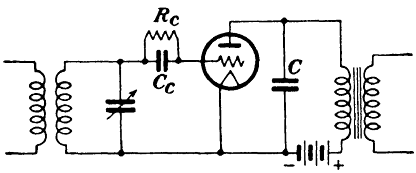


FIG. 224.—Circuit for grid-leak detection.

grid leak and condenser. The principle of plate rectification is commonly used in vacuum-tube voltmeters, which will be discussed later in the chapter.

142. Grid-current Detection.—This mode of operation, often called grid-leak detection, depends upon the curvature of the I_g-E_g characteristic of the tube. The circuit is shown in Fig. 224. The grid return is connected to the positive end of the filament which causes grid current to flow. In the case of equipotential cathodes the grid return is connected directly to the cathode. Grid current also flows in this case, even though the grid is at the same potential as the cathode, owing to the initial velocity of emission of the electrons from the cathode.

The I_g-E_g characteristic of a typical heater-type tube is shown in Fig. 225, and it will be observed that a negative bias of approximately 1 volt would be required to reduce the grid current to zero. This grid current flows through the high resistance R_c and produces a voltage drop across the resistor which imposes a small negative bias upon the tube. In the example shown a grid leak of 200,000 ohms produces a negative bias of 0.7 volt, which becomes the operating potential E_c of the grid when no signal is

impressed. This value of R_c is only about one-tenth of what would be actually used and is for purposes of illustration only. The operating point is determined by the intersection with the characteristic of a line drawn from $E_g = 0$ which makes an angle with the vertical whose tangent is R_c, taking into account the differences in the current and voltage scales.

When a modulated signal is impressed, rectification takes place because of the curvature in the grid-current characteristic. This

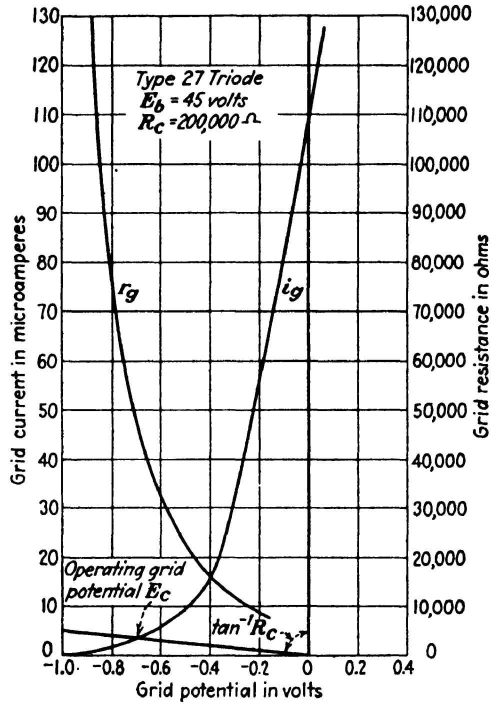


FIG. 225.—Grid-current and grid-resistance curves of a typical triode having an equipotential cathode.

causes the average grid current to vary in exactly the same manner as did the average plate current i_s in Fig. 222 in the case of the plate rectifier. This audio-frequency variation in the average value of the grid current is made to flow through R_c, as the reactance of C_c (usually about 250 $\mu\mu$f) is too high to by-pass any appreciable portion. The radio-frequency portion of the grid current passes through C_c with negligible voltage drop. The audio-frequency component of grid current flowing through

R_c produces a voltage drop across this resistance which causes the average grid potential to vary in the manner shown in (*b*) of Fig. 226, producing the corresponding variations in the plate current shown in (*c*). In other words, the audio-frequency voltage developed across R_c, since it exists between the grid and filament, is amplified by the detector tube functioning in the manner of an ordinary audio-frequency amplifier. It will be noted that

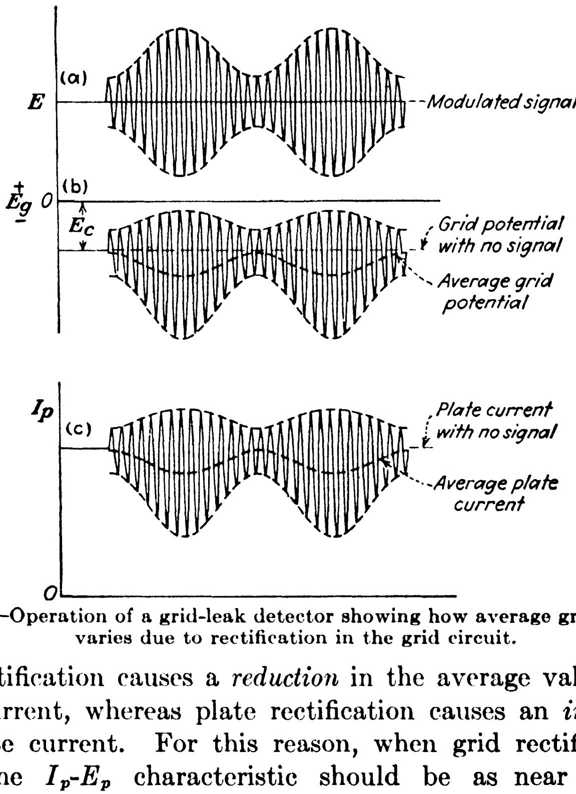


FIG. 226.—Operation of a grid-leak detector showing how average grid potential varies due to rectification in the grid circuit.

grid rectification causes a *reduction* in the average value of the plate current, whereas plate rectification causes an *increase* in the plate current. For this reason, when grid rectification is used, the I_p-E_p characteristic should be as near linear as possible so as to avoid plate rectification, which tends to offset the effects of grid detection.

As in the case of the plate current, the grid current can also be represented by a series of the form

$$I_g = I_0 + c_1e_g + c_2e_g^2 + \cdots \tag{10}$$

where the squared term is responsible for the rectification. If the signal voltage is small, the terms of higher order may be neglected, as before.

Evaluating the constant c_2 by the method used in (24), Chap. VI, this term becomes

$$i_g = \frac{1}{2}\frac{\partial^2 i_g}{\partial e_g^2}e_g^2 \tag{11}$$

If a modulated signal of the form

$$e_g = E(1 + m \sin \omega_s t) \sin \omega_c t \tag{12}$$

is impressed on the grid, (11) becomes

$$i_g = \frac{1}{2}\frac{\partial^2 i_g}{\partial e_g^2}E^2(1 + m \sin \omega_s t)^2 \sin^2 \omega_c t \tag{13}$$

The terms of audio frequency in (13) will be

$$i_s = \frac{1}{2}\frac{\partial^2 i_g}{\partial e_g^2}E^2\left(m \sin \omega_s t - \frac{m^2}{4} \cos 2\omega_s t\right) \tag{14}$$

which is similar to (5). Distortion in the form of 25 per cent second harmonic will again be produced with a completely modulated signal, which is characteristic of all square-law detectors.

The coefficient $\partial^2 i_g/\partial e_g^2$ is to be evaluated at the average direct-current potential of the grid, or at point E_c in Fig. 225. Evidently

$$\frac{\partial i_g}{\partial e_g} = \frac{1}{r_g} \tag{15}$$

where r_g is the dynamic or alternating-current resistance between the grid and filament terminals of the tube. Like r_p, it varies with the direct-current potential of the grid, as shown in Fig. 225. Its values may be determined graphically by drawing tangents to the curve of grid current, or preferably, by means of a bridge measurement similar to the circuit used in the measurement of r_p in Fig. 89, page 141.

The second derivative is

$$\frac{\partial^2 i_g}{\partial e_g^2} = \frac{\partial \frac{1}{r_g}}{\partial e_g^2} = -\frac{1}{r_g^2}\frac{\partial r_g}{\partial e_g} \tag{16}$$

which may be evaluated by changing the grid potential by a small increment Δe_g above, and then below, the operating potential E_c; the corresponding values of the grid resistance being measured for each increment by means of a bridge. Calling these values of grid resistance r_g' and r_g'', then

$$\frac{r_g' - r_g''}{2\Delta e_g} = \frac{\partial r_g}{\partial e_g} \tag{17}$$

to a good degree of approximation. The value of Δe_g used in the measurement should be in the vicinity of 0.05 volt. Since r_g has a negative slope, (17) will also be negative.

It has been shown by F. E. Terman[2] that the detection constant v, defined as

$$v = \frac{2r_g}{\partial r_g/\partial e_g} \tag{18}$$

is practically constant for a given tube over a wide range of operating values of r_g. Values of this detection constant for

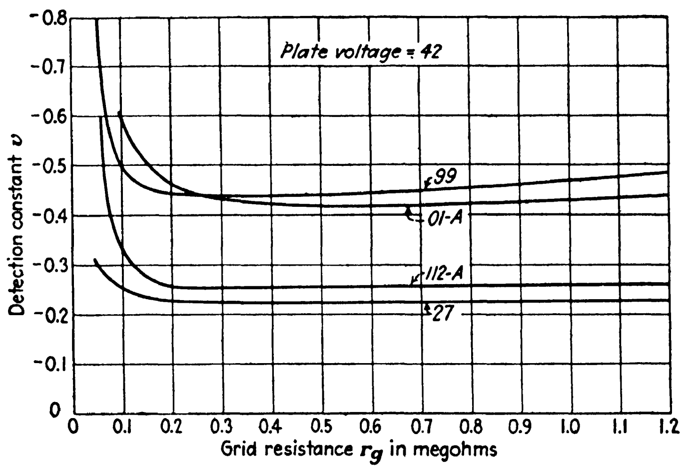


FIG. 227.—Variation of the detection constant v with grid resistance for several typical detector tubes.

several typical detector tubes given in Terman's paper are shown in Fig. 227. The value of grid resistance r_g will depend upon the bias voltage E_c at the operating point, which, in turn, is governed by the resistance R_c of the grid leak used, as illustrated in Fig. 225. It is interesting to note that the sensitivity of the tubes using oxide-coated cathodes is approximately twice as great as of those using thoriated tungsten as an emitter.

[2] Some Principles of Grid-leak Grid Condenser Detection, *Proc. I.R.E.*, vol. 16, p. 1384, October, 1928; also, F. E. TERMAN and T. M. GOOGIN, Detection Characteristics of Three-element Vacuum Tubes, *Proc. I.R.E.*, vol. 17, p. 149, January, 1929.

Substituting (16) and (18) in (14), the audio-frequency current flowing in the grid circuit will be

$$i_s = \frac{E^2}{vr_g}\left(m \sin \omega_s t - \frac{m^2}{4} \cos 2\omega_s t\right) \qquad (19)$$

where E is the maximum value of the unmodulated carrier applied to the grid of the tube. This will be approximately the same as the voltage developed across the tuned circuit of Fig. 224 if the impedance drop across the grid condenser C_c is negligible at radio frequencies.

This rectified current i_s may be thought of as having been produced by a fictitious generator E_r acting in series with r_g, as shown in the equivalent circuit of Fig. 228. This proposition of equivalence was first proved by J. R. Carson with reference to plate-circuit modulation.[3] There would be one such generator for each frequency component of i_s, and each frequency term can be considered individually without regard for the currents or voltage drops produced by the others. The tuned input is neglected in the equivalent circuit as the coil offers negligible impedance to currents of audio frequency. The condenser C_g represents the tube-input capacitance and is dependent upon the interelectrode capacitances and the nature of the load impedance in the plate circuit, as pointed out in Chap. VIII. It will have one value for the audio frequencies and a much lower value C'_g for the impressed radio-frequency signal, owing to the effect of the by-pass condenser C across the load in the plate circuit.

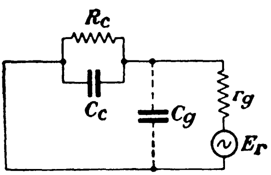


FIG. 228.—Equivalent grid circuit of a grid-leak detector.

The value of the rectified voltage E_r will be

$$E_r = i_s r_g = \frac{E^2}{v}\left(m \sin \omega_s t - \frac{m^2}{4} \cos 2\omega_s t\right) \qquad (20)$$

From Fig. 228, a portion of this voltage will appear across the impedance Z_c of the grid leak and grid condenser in parallel, the latter being augmented by value of C_g. This voltage will be

[3] The Equivalent Circuit of the Vacuum-tube Modulator, *Proc. I.R.E.*, vol. 9, p. 243, June, 1921.

$$E_s = \frac{Z_c}{r_g + Z_c} E_r \tag{21}$$

where

$$Z_c = \frac{R_c}{1 + j\omega_s R_c(C_c + C_g)} \tag{22}$$

The audio-frequency voltage E_s across the grid leak will then be amplified by the tube functioning as an audio-frequency amplifier. The value of E_s, omitting the second-harmonic-distortion term, will be

$$E_s = \frac{Z_c}{r_g + Z_c} \times \frac{E^2}{v} m \sin \omega_s t \tag{23}$$

The audio-frequency output voltage of the detector will be

$$E_d = i_p Z_b = \frac{\mu E_s}{r_p + Z_b} Z_b \tag{24}$$

where Z_b is the vector expression for the plate-load impedance in the output circuit of the detector at the radian frequency ω_s.

The amplitude E of the unmodulated carrier voltage in (23) is the actual value impressed across the grid-filament terminals of the tube. This value will be slightly less than the voltage E_o across the tuned input circuit by the amount of the drop across Z_c at radio frequencies. The relation between these two is approximately

$$E = E_o \frac{C_c}{C_c + C_g'} \tag{25}$$

under the assumption that E_o divides inversely as the capacitances of C_c and C_g', the latter being the input capacitance of the tube at radio frequencies. This expression neglects the fact that these capacitances are shunted by R_c and r_g, respectively, which may be taken into account in (25) by substituting the vector expressions Z_c and Z_g for C_c and C_g', if necessary.

143. Frequency Distortion in Grid-leak Detectors.—The audio-frequency voltage E_s developed across the grid-leak and condenser combination Z_c tends to fall off at the higher values of audio frequency owing to the shunting effect of C_c. This may cause serious frequency distortion, particularly if a large value of grid-leak resistance is used in an attempt to secure high sensitivity.

The effect of grid-leak resistance is shown in Fig. 229, taken from Terman's paper,[2] for a typical detector tube operating at a plate potential of 22 volts. Reducing the value of R_c will lessen the frequency distortion, but it will also impair the sensitivity. Keeping R_c fixed while the size of the grid condenser C_c is reduced will produce similar effects. Consequently, the choice of these

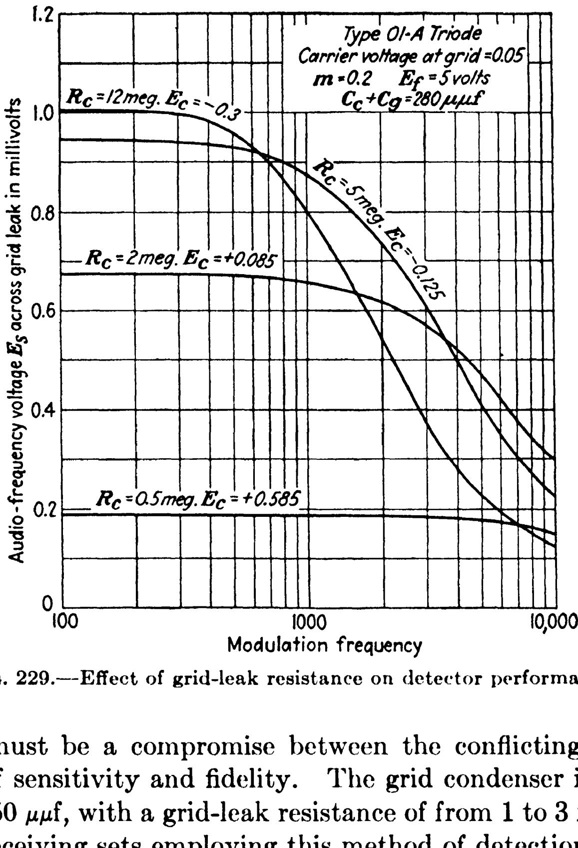


FIG. 229.—Effect of grid-leak resistance on detector performance.

values must be a compromise between the conflicting requirements of sensitivity and fidelity. The grid condenser is usually about 250 $\mu\mu$f, with a grid-leak resistance of from 1 to 3 megohms in the receiving sets employing this method of detection.

Further discrimination against the higher audio frequencies is caused by the plate by-pass condenser C in Fig. 224, which provides a path of low impedance for the radio-frequency components of the plate current. Since the primary impedance of the audio-frequency transformer is quite large this condenser may also by-pass a portion of the higher audio-frequency com-

[2] *Loc. cit.*

ponents. This possible cause of frequency distortion is present in both grid and plate detectors.

The detector is frequently coupled to the audio-frequency-output device by a low-pass filter section, as shown in Fig. 230. A structure of this type composed of resistanceless elements, when properly terminated, will transmit without attenuation all frequencies up to the cut-off frequency f_c, which is given by[4]

$$f_c = \frac{1}{\pi\sqrt{LC}} \tag{26}$$

where

$$L = \frac{R}{\pi f_c} \tag{27}$$

$$C = \frac{1}{\pi f_c R} \tag{28}$$

and R is the value of the terminating impedance at each end of the filter. From the standpoint of filter theory the structure should be connected between a source having an internal resistance R and a load of the same value. This requirement is unattainable when an audio-frequency transformer is used and the filter must therefore be designed for a load R equal to r_p of the detector tube under operating conditions, and consequently will be properly terminated at one end only. This necessary compromise, however, will not usually alter the performance of the filter very seriously.

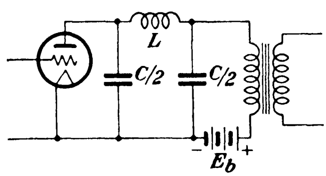


FIG. 230. —Low-pass filter section in the plate circuit of a detector.

In addition to transmitting the desired range of audio frequencies without discrimination, the filter element effectively excludes radio-frequency currents from the audio-frequency amplifier where they may be a potential source of trouble and cause feed-back.

Another disadvantage of the grid-leak detector is that the tuned circuit is shunted in effect by the grid resistance r_g which is often low enough to cause an appreciable reduction in the gain and selectivity of the preceding radio-frequency amplifier stage.

[4] K. S. JOHNSON, "Transmission Circuits for Telephonic Communication," Chap. XVII, D. Van Nostrand Company, Inc., 1925.

144. Grid-leak Power Detectors.[5]—The grid-leak detector just considered is not capable of furnishing a very large audio-frequency output voltage without serious distortion due to overload. However, if the grid leak and condenser are both reduced in size and the plate voltage is materially increased, signal voltages of 5 volts or more may be impressed on the detector and with

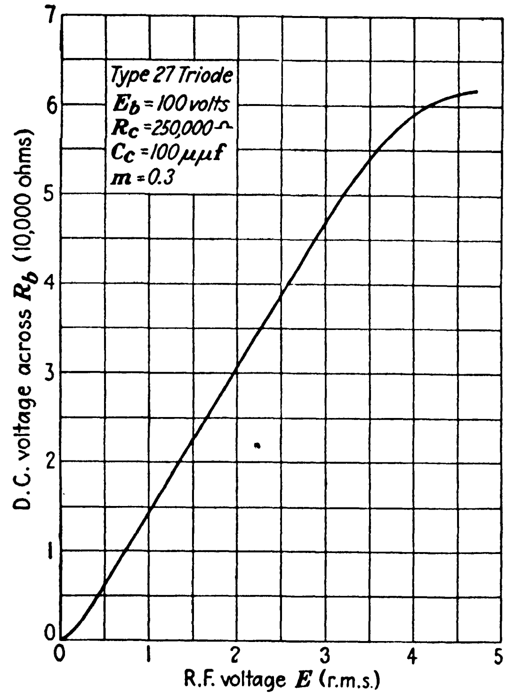


FIG. 231.—Characteristics of a grid-leak power detector showing the approximately linear relation between impressed high-frequency carrier voltage and audio-frequency output voltage.

a corresponding increase in the audio-frequency output. The percentage of distortion will be less than it formerly was with small-signal voltages and the output will be a linear function of the impressed signal. At very small signal voltages square-law operation again results. The maximum signal that can be handled is limited by the distortion due to the increasing amount of plate rectification. This prevents the rectified output of the

[5] F. E. TERMAN and N. R. MORGAN, Some Principles of Grid-leak Power Detection, *Proc. I.R.E.*, vol. 18, p. 2160, December, 1930.

detector from continuing to increase directly with the impressed signal voltage, as illustrated in Fig. 231.

The principal difference in the mode of operation between these two types of grid-leak detectors is that in the power type the grid potential swings negative by an amount sufficient to stop the flow of grid current for the greater portion of the negative half cycle. During the positive half cycle of the carrier voltage an impulse of grid current flows which charges the grid condenser negatively. Some of this accumulated charge, representing a negative bias on the grid, escapes through the grid leak during the next negative half cycle and is replenished during the following positive swing. Grid current will therefore flow only during a brief positive portion of the cycle. This mode of operation differs from the small-signal type of detector, in that in the latter grid current flows continuously, since the grid is never driven to a sufficiently negative potential to reduce i_g to zero. The charge on the grid condenser under these conditions leaks off through r_g which is much lower in value than the grid-leak resistance R_c. In the large-signal detector, r_g is infinite during the greater portion of the cycle. The operation is the same as in Fig. 226, except that the amplitude of grid voltage in (*b*) is much greater and the cut-off point of grid current is approximately E_c. In the small-signal case, E_c is less than cut-off, as shown in Fig. 225.

In order for the average grid potential to follow the audio-frequency variations in the envelope of the carrier wave, the time constant of the circuit composed of the grid leak and grid condenser must be properly chosen so as to enable the grid condenser to charge and discharge at least as rapidly as the signal amplitude is changing. This requires smaller values of grid-condenser capacitance and grid-leak resistance than are used in the small-signal type of detector. Average values for these items in a power detector are a grid-leak resistance of $\frac{1}{4}$ megohm and a 100-$\mu\mu$f grid condenser. The effective capacitance of the latter is augmented by C_g of the tube.

The rate of change in the envelope of the carrier is a function of the modulating frequency f_s and the degree of modulation m. It can be shown that the distortion caused by the grid-leak and condenser will be small if these items satisfy the relation[5]

[5] *Loc. cit.*

$$\frac{X}{R_c} \geqq \frac{m}{\sqrt{1 - m^2}} \tag{29}$$

where

$$X = \frac{1}{\omega_s(C_c + C_g)} \tag{30}$$

The variations in the average negative potential of the grid are amplified in the ordinary manner by the detector tube. The magnitude of these variations will depend upon the rectification efficiency β which will usually be from 60 to 85 per cent, depending on the type of cathode in the tube. Thus the audio-frequency signal voltage E_s that in effect is applied to the grid of the detector tube is

$$E_s = \beta m E \tag{31}$$

where E is the maximum amplitude of the unmodulated carrier voltage applied to the detector. The audio-frequency output voltage E_d can be obtained from (24).

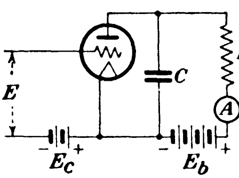


FIG. 232.—Circuit for determining performance of grid-bias detectors.

145. Grid-bias Detectors.—This method of detection, also known as linear plate detection, makes use of a tube which is biased practically to cut-off so that the plate current is nearly zero when no signal is applied to the grid. The operation is almost identical with that of the Class B amplifier shown in Fig. 184, except that the signal should not be allowed to swing the grid positive unless the amplifier supplying the signal has good voltage regulation and is capable of furnishing the necessary power represented by the flow of grid current. This means that the signal will, in general, be confined to that portion of the dynamic characteristic lying to the left of $E_g = 0$. It is therefore more difficult with this type of detection to secure the same degree of linear operation possible with Class B amplifiers, which utilize a much larger portion of the characteristic and thereby reduce the effect of the appreciable curvature in the vicinity of cut-off.

The performance of a grid-bias detector can be readily determined experimentally by means of the circuit of Fig. 232.[6] An alternating voltage E of any convenient frequency such as 60

[6] STUART BALLANTINE, Detection at High Signal Voltages, *Proc. I.R.E.*, vol. 17, p. 1153, July, 1929.

cycles is applied to the grid of the tube and the average value of plate current I_b is read by a suitable direct-current milliammeter. This current will increase from I_b to I'_b when E is impressed. The by-pass condenser C should be large enough so as to offer a path of low impedance to the alternating component of plate

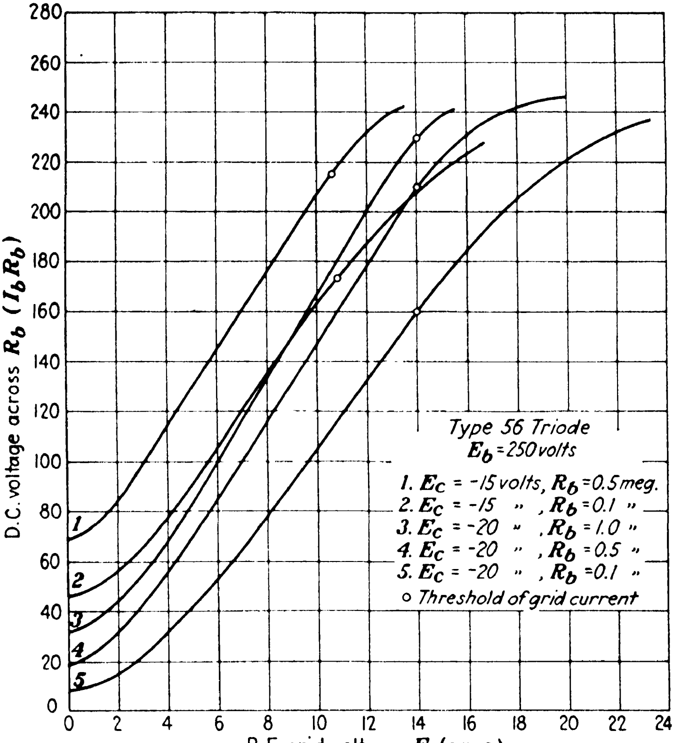


FIG. 233.—Performance curves of a triode as a grid-bias detector.

current. If E is thought of as rising and falling in amplitude by an amount $E(1 \pm m)$, where m is the degree of modulation, its action would be similar to that of a modulated signal. The current through R_b would then also increase and decrease from I'_b to I''_b. The useful output of the detector across R_b at the variation frequency (which corresponds to ω_s in this case) will be

$$E_d = \frac{I'_b - I''_b}{2} R_b \tag{32}$$

This will be the *maximum value* of the audio-frequency voltage.

Performance curves obtained in the above manner are shown in Fig. 233 for a typical triode under various conditions of operation. Figure 234 shows similar curves for various types of tubes commonly used as grid-bias detectors. The slopes of these curves are directly proportional to sensitivities of the tubes as detectors. The increased sensitivity of the pentode and tetrode types is

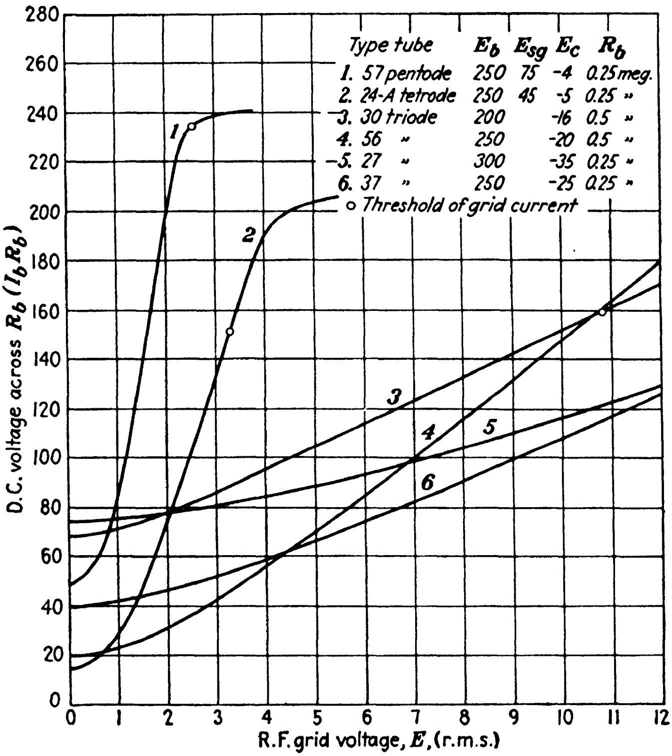


FIG. 234.—Performance curves of various grid-bias detectors.

apparent from curves 1 and 2 in Fig. 234. Their distortion is usually somewhat greater on high degrees of modulation, however, than with triodes. The sensitivity of the detector is increased as R_b is made larger, but with an increased loss of the higher modulation frequencies due to the decreased negligibility of the reactance of condenser C at these frequencies. The maximum value of R_b that can be used is therefore limited by the amount of frequency distortion that can be tolerated.

All of the detection characteristics show appreciable curvature for small values of impressed signal and the operation will therefore follow a square law when E is small. When this type of detection is used in broadcast receivers, provision should be made to impress a sufficiently large value of signal voltage on the grid in order to avoid the distortion pertaining to square-law operation. This is one of the advantages of automatic volume control in receiving sets.

When the applied signal voltage is sufficiently large, the modulated envelope is impressed upon the straight portion of the

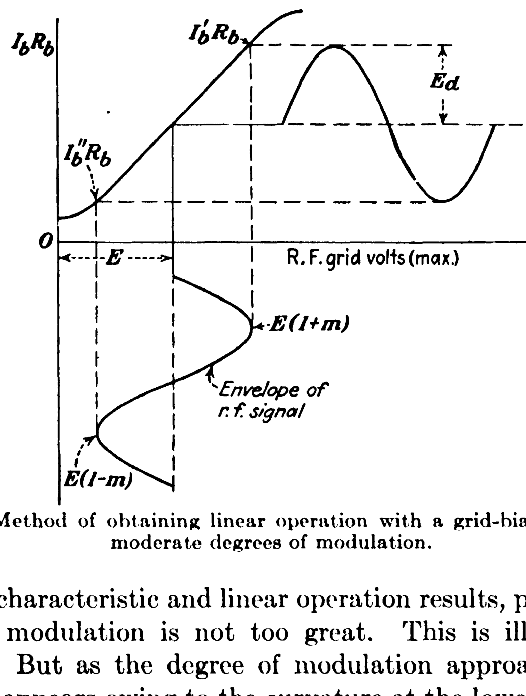


FIG. 235.—Method of obtaining linear operation with a grid-bias detector for moderate degrees of modulation.

detection characteristic and linear operation results, provided the degree of modulation is not too great. This is illustrated in Fig. 235. But as the degree of modulation approaches unity, distortion appears owing to the curvature at the lower end of the characteristic. Grid-bias detectors will therefore only produce linear operation when the signal voltage applied is sufficiently large and the degree of modulation is not too great.

Detection characteristics of the type shown in Figs. 233 and 234 can be obtained for any value of load resistance R_b by means of the transrectification diagram of Fig. 236. These curves show the relation between the direct-current values of plate current and plate voltage with various radio-frequency voltages applied

to the grid, and may be determined experimentally by means of Fig. 232, by making R_b zero. The alternating voltage E impressed on the grid can be of low frequency. By drawing a

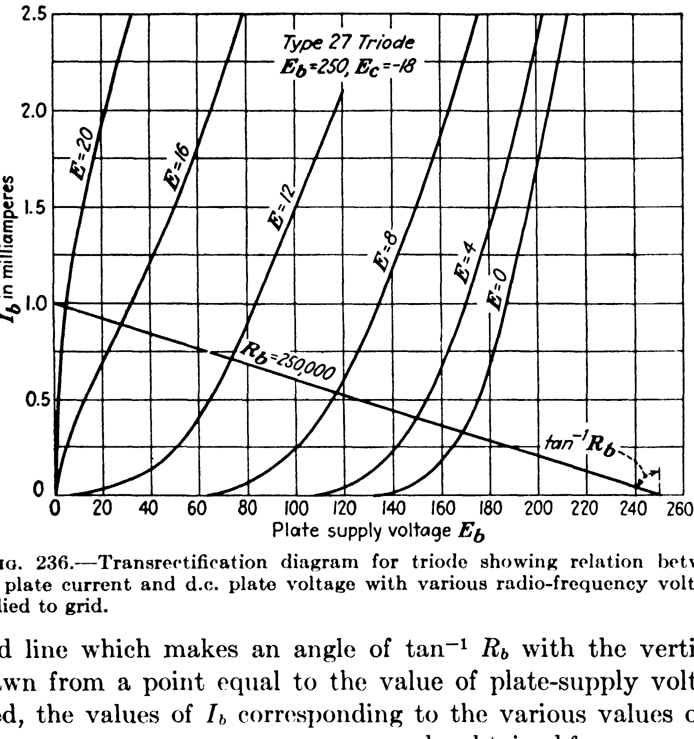


FIG. 236.—Transrectification diagram for triode showing relation between d.c. plate current and d.c. plate voltage with various radio-frequency voltages applied to grid.

load line which makes an angle of $\tan^{-1} R_b$ with the vertical, drawn from a point equal to the value of plate-supply voltage used, the values of I_b corresponding to the various values of E may be obtained for any particular value of R_b. Knowing I_b, the voltage drop across R_b may be computed. Thus in Fig. 236, with an unmodulated carrier of 8 volts, the d-c plate current will be 0.55 ma. When the signal is completely modulated, the carrier amplitude will vary from zero to 16 volts at an audio rate. The plate current will vary 0.3 to 0.8 ma (approximately), and the plate voltage will vary from 168 to 28 volts. The approximate r.m.s. audio voltage developed across the 250,000-ohm plate load will be about 50 volts.

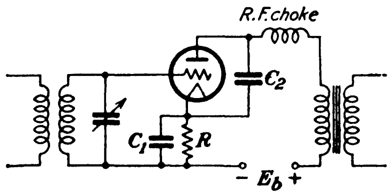


FIG. 237.—Grid-bias detector circuit self-biased by means of cathode resistor R.

The audio-frequency amplifier following the detector is connected across R_b by means of a grid-leak and blocking condenser, the details of which have already been discussed in Sec. 62, Chap. VII. This reduces the effective resistance of the load in the plate circuit of the detector for audio-frequency currents from R_b to $\frac{R_bR_c}{R_b + R_c}$, where R_c is the grid-leak resistance of the following amplifier stage, assuming the reactance of the blocking condenser to be negligible. The construction of the load line in Fig. 236 will then be similar to Fig. 108. When a reactance load, such as the primary of an audio-frequency transformer, is used in place of a resistance the load line becomes an ellipse, as discussed on page 162 in connection with Fig. 109.

In battery-operated receiving sets the required biasing voltage can be readily obtained from a suitable C battery. When rectified alternating current is used as a source of B supply, the necessary grid bias can be obtained by means of a resistor R in series with the cathode, as shown in Fig. 237. The by-pass condenser C_1 across this resistor is not required to be so large as in the case of the self-biased audio-frequency amplifiers discussed in Sec. 76. This is due to the much larger value of resistance needed in order to bias the tube to almost cut-off, so that only a moderate value of C_1 is sufficient to avoid degenerative effects at low audio frequencies.

Grid-bias detectors do not ordinarily draw grid current and therefore do not impose a load upon the tuned input circuit as is the case with grid-leak detection. However, they are less sensitive than grid-leak detectors.

146. Diode Detectors.—In recent years, with the advent of automatic volume control in receiving sets, diode detectors have come into extensive use. While not nearly so sensitive as some of the types already discussed, since there is no amplifying action, diode detector circuits are comparatively free from distortion. This is due to the relatively low internal resistance of the diode in the conducting direction, which, while variable, can be made small in comparison to the external load resistance, resulting in approximately linear operation.

A typical diode-detector circuit is shown in Fig. 238. The value of R_b is ordinarily from 0.5 to 1 megohm, while the by-pass condenser C is in the vicinity of 150 $\mu\mu$f at broadcast frequencies.

In superheterodyne receivers, where the impressed carrier frequency is much lower, a condenser of two to three times this capacitance should be used. These values of R_b are large compared with the internal resistance of the tube, except at small values of plate current, so that the rectified voltage E_d across the load resistance R_b is nearly equal to the maximum value of the alternating voltage E across the tuned circuit. This assumes that the condenser C offers negligible reactance to the frequency of E, which is not always true in practice.

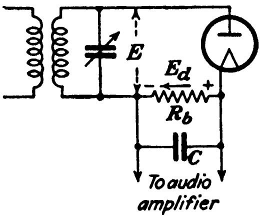


Fig. 238.—Diode detector circuit.

The performance of a diode as a detector is shown in Fig. 239. The experimental data for such curves can be obtained by the

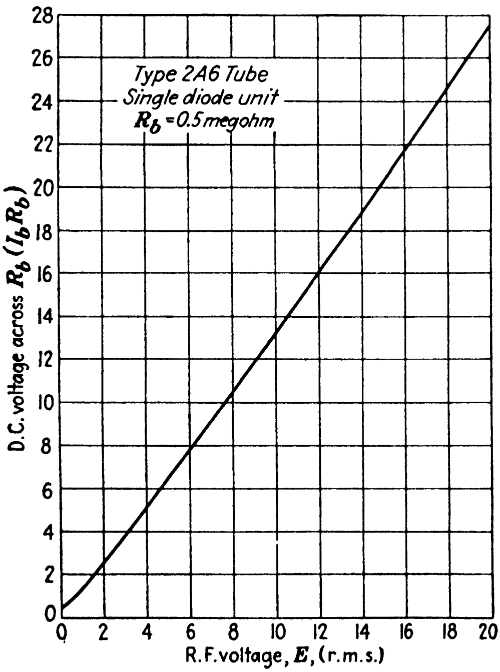


Fig. 239.—Performance curve of a typical diode detector.

circuit of Fig. 238 by impressing an alternating voltage E of any convenient frequency on the detector and observing the corre-

sponding values of rectified current I_b flowing through R_b by means of a suitable microammeter. The output voltage E_d of the detector will be I_bR_b. If the impressed frequency is low the by-pass condenser C should be about 2 μf. Ordinary triodes may be used as diode detectors by tying the grid and plate together, or by connecting the plate to the cathode as shown in Fig. 240. The performance is virtually the same with either type of connection. The arrangement in (b) permits the plate to serve as an electrostatic shield enclosing the grid, as the cathode to which it is connected is usually at ground potential. A triode connected for diode operation will have practically the same type of characteristic as Fig. 239. As long as the internal resistance

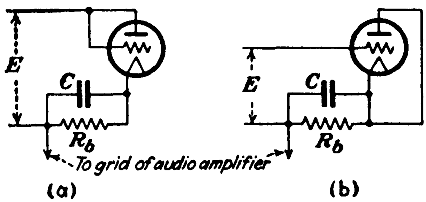


FIG. 240.—Triodes connected for diode operation as detectors.

of the tube used is small in the conducting direction compared with the load resistance R_b, the rectified output voltage E_d of the detector will be nearly equal to the maximum value of the impressed alternating voltage E, and consequently there will be no material differences in the sensitivities of various types of diodes.

Tubes having two diode elements are available. These duplex tubes also have triode or pentode elements incorporated in the same bulb, a single cathode sleeve of the heater type being common to all. One of the two diode elements may be used for purposes of automatic volume control, as described in Chap. XIII, or the two may be connected in parallel as a half-wave rectifier with a corresponding decrease in the internal resistance. Another possible combination is to connect the two diode units as a full-wave rectifier. The latter arrangement is shown in Fig. 241, together with the connections for the included triode unit as a resistance-coupled audio-frequency amplifier. Only half of the available signal voltage E developed across the tuned input is impressed on each diode so that the corresponding output voltage E_d across R_{b1} will likewise be cut in half. This disad-

vantage is partially offset by the theoretical absence of carrier-frequency voltage across R_{b1} if the two diodes are exactly alike, which greatly reduces the amount of carrier-frequency filtering required. With half-wave rectification provision in the form of an adequate filter must be made to prevent radio-frequency voltages from reaching the grid of the following audio-frequency amplifier.

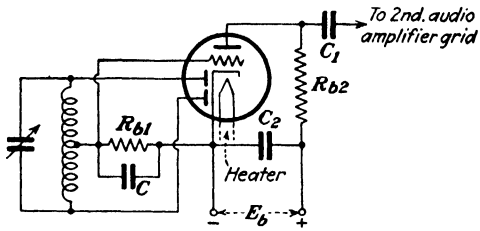


FIG. 241.—Full-wave detector circuit using a duplex-diode triode tube. The triode unit serves as the first stage of audio-frequency amplification.

The grid bias for the triode unit is obtained from the voltage drop across R_{b1}. As this voltage fluctuates with the intensity of the received signal, the bias will likewise fluctuate, and will fall to zero when there is no radio-frequency input. Excessive triode plate current may therefore result unless limited by sufficient resistance in series with plate-supply voltage. When resistance coupling is used, the coupling resistor R_{b2} will assist in limiting the current. With very strong signals the triode may be biased beyond cut-off so that this method of "diode biasing" is ordinarily practical only when automatic volume control is

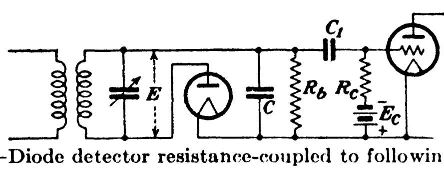


FIG. 242.—Diode detector resistance-coupled to following amplifier.

employed, which will insure a reasonably constant value of applied radio-frequency voltage. A fixed bias is usually preferable, particularly with a high-mu triode. In order to impress a fixed bias on the grid of the following amplifier without at the same time imposing a bias on the diode a coupling condenser and grid resistor are required, as shown in Fig. 242. This introduces complications when the rectification of a modulated signal is

considered, as the load in series with the diode will have a value of R_b for direct current and a lower value made up of R_b and R_c in parallel (assuming C_1 large) for the modulating frequency.

This may be seen from the rectification diagram of the diode, which can be obtained experimentally by means of the circuit of Fig. 243. The alternating voltage E is held constant while the bias voltage E_c (which simulates the drop across R_b) is varied, the rectified current I_b being read by the microammeter A. The rectification characteristics obtained in this manner are shown in Fig. 244 for various values of E. The performance with various values of load resistance may be obtained by drawing a line through the origin which makes an angle of $\tan^{-1} R_b$ with the vertical. This construction assumes that R_b is by-passed so as to offer negligible impedance to the impressed radio-frequency signal E. But when the amplitude of E varies at an audio-frequency rate owing to modulation, the impedance offered to modulating frequency will be $R'_b = \dfrac{R_bR_c}{R_b + R_c}$, assuming the react-

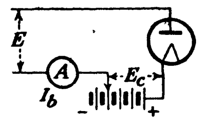


FIG. 243.—Circuit for experimental determination of diode rectification diagram.

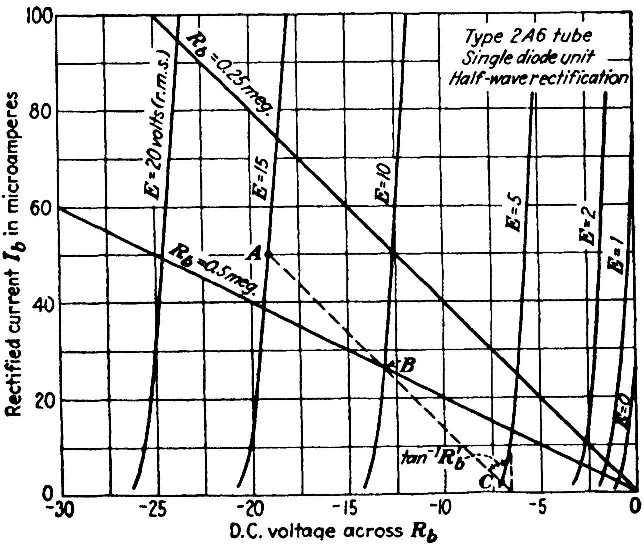


FIG. 244.—Rectification diagram for a diode showing relation between d-c current and output voltage with various radio-frequency voltages applied.

ance of the coupling condenser C_1 in Fig. 242 to be negligible. Thus, if R_b and R_c are both 0.5 megohm and the effective value of the unmodulated signal is 10 volts, the path of operation will be along the dotted line ABC. If 50 per cent modulation is assumed, E will vary from 15 to 5 volts, causing point C to be nearly at cut-off, and with greater percentages of modulation, distortion will result, as I_b cannot flow in the reverse direction. Therefore, in order to prevent distortion at high values of modulation, R_b and R_b' should be made nearly equal in value, which means that R_c should be large compared to R_b. From this it is evident that diode detectors also produce distortion when the degree of modulation is high.[7] This difficulty is absent in the circuit of Fig. 241, since $R_b' = R_{b1}$ in this case.

The output energy of the diode detector must be supplied by the preceding radio-frequency amplifier, and hence the gain and sharpness of tuning of this stage are affected to some degree, although usually to a much lesser extent than with grid-leak detection. The input resistance of the diode detector is essentially constant and independent of the impressed voltage E for load resistances large compared to the internal tube resistance. Under this condition the value of input resistance is approximately $\frac{1}{2}R_b'$.

147. Heterodyne Detection.—When two slightly different frequencies are superimposed, the envelope of the resultant wave will vary in amplitude at a rate equal to the frequency difference between the two. In the case of two tuning forks this difference frequency gives rise to the phenomenon of "beats" when the frequencies of the two are not exactly the same. The appearance of the resultant wave is shown in Fig. 245. When the two components are of equal amplitude, as in (*c*), the shape of the envelope is similar to the output of a full-wave rectifier. Unequal amplitudes of the two components cause pulsations in the amplitude of the resultant wave equal to the double amplitude of the smaller of the two components, as in (*d*).

If the resultant wave is rectified, the average value of the rectified current will vary at the beat-frequency rate. The process of impressing two voltages differing in frequency upon a detector and extracting the beat frequency is called *heterodyne*

[7] C. E. KILGOUR and J. M. GLESSNER, Diode Detection Analysis, *Proc. I.R.E.*, vol. 21, p. 930, July, 1933.

detection and has important applications in radio communication. The received dots and dashes in continuous-wave telegraphy are heterodyned with a local oscillator differing in frequency from that of the received signals by some audible frequency such as 1000 cycles. The rectified output of the detector will contain the desired beat frequency which can be used to actuate a telephone receiver. In the superheterodyne system of reception, so widely used in modern receiving sets, the incoming modulated signals are superimposed on the output of a local oscillator having a frequency higher or lower than the received carrier by a desired amount. The resultant beat or intermediate frequency can thus be made much lower than the received frequency for convenience of amplification. The amplitude of the beats will vary in accordance with the modulated envelope of the incoming frequency and the process of heterodyne detection has in effect merely reduced the carrier frequency from its former value to a new lower value. The amplifier following the detector can then be initially tuned to a fixed value of beat frequency. All incoming signals are converted to the same beat frequency by changing the frequency of the local oscillator. The difficulty of getting the various tuned circuits of the amplifier to "track" properly is thereby avoided since their tuning adjustments remain fixed. In other words, the carrier frequency of the signal is changed to suit the amplifier, rather than adjusting the amplifier to fit the signal.

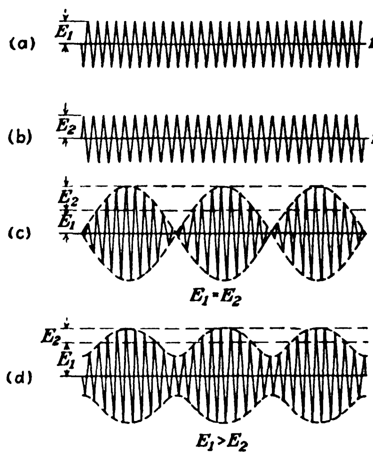


FIG. 245.—Result of combining two waves of slightly different frequencies, which causes the resultant wave to pulsate in amplitude at a rate equal to the frequency difference of the two component waves.

The shape of the resulting envelope in Fig. 245*c* when E_1 and E_2 are of equal amplitude is such that when impressed upon a square-law detector, the beat-frequency output will be sinusoidal and no harmonics of the beat frequency will be produced. When impressed upon a linear detector, the exact shape of the applied

envelope will be reproduced in the output and beat-frequency-distortion terms will be present. When $E_1 = E_2 = E$, the output wave will be identical with that of a full-wave rectifier, and may be represented by a Fourier series of the form

$$e = E(1.274 - 0.8488 \cos \omega t - 0.1697 \cos 2\omega t - 0.0728 \cos 3\omega t - 0.0404 \cos 4\omega t - \cdots) \quad (33)$$

where $\omega = 2\pi(f_1 - f_2)$, f_1 and f_2 being the frequencies of the original waves. When one of the superimposed voltages is very small compared to the other, the amplitude of the pulsations becomes very small and the envelope of Fig. 245*d* becomes nearly sinusoidal. This is the usual condition in practice, since the output of the local oscillator is large compared to the voltage of the received signal. Either a square-law or a linear detector can be used as the effective modulation under these circumstances is small and the beat-frequency-distortion terms will also be small. The original speech-modulated signal, whether completely modulated or not, will have very little influence upon the degree of effective modulation when E_1 and E_2 differ considerably in size, and consequently the audio frequencies will suffer little distortion, even if the heterodyne detector follows a square law.

148. Regenerative Detectors.—It has been shown in Chap. VII that if a portion of the output energy of an amplifier is fed back into the input circuit in the proper phase, the amplification can be greatly increased. The preceding types of detectors all have appreciable components of carrier frequency in their plate circuits and it is therefore possible to obtain regeneration by the use of suitable feed-back circuits in all of these types, with the exception of the diode, where amplifying action is absent. The sensitivity of the detector is greatly increased by regeneration, which introduces negative resistance into the tuned input circuit. This principle was extensively used in the earlier broadcast-receiving sets and it is still used in receivers designed primarily for the reception of telegraphic code signals, especially on short waves. By increasing the feed-back to the point where the detector breaks into oscillation the tube may be made to function simultaneously as a local oscillator as well as a detector; this heterodyne action enabling continuous-wave signals to be made audible. This is known as *autodyne reception*.

The most commonly used regenerative detector circuits are illustrated in Fig. 246. The circuit shown in (*a*) is identical with the tuned-grid-oscillator circuit and the control of regeneration is accomplished by varying the mutual inductance M between L_1 and L_2, the latter being usually called a "tickler" coil. As M is gradually increased, the sensitivity increases and the circuit finally breaks into oscillation. The arrangement of Fig. 246*b* makes use of the negative input resistance of a tube when the plate circuit contains an inductive reactance. The variation of

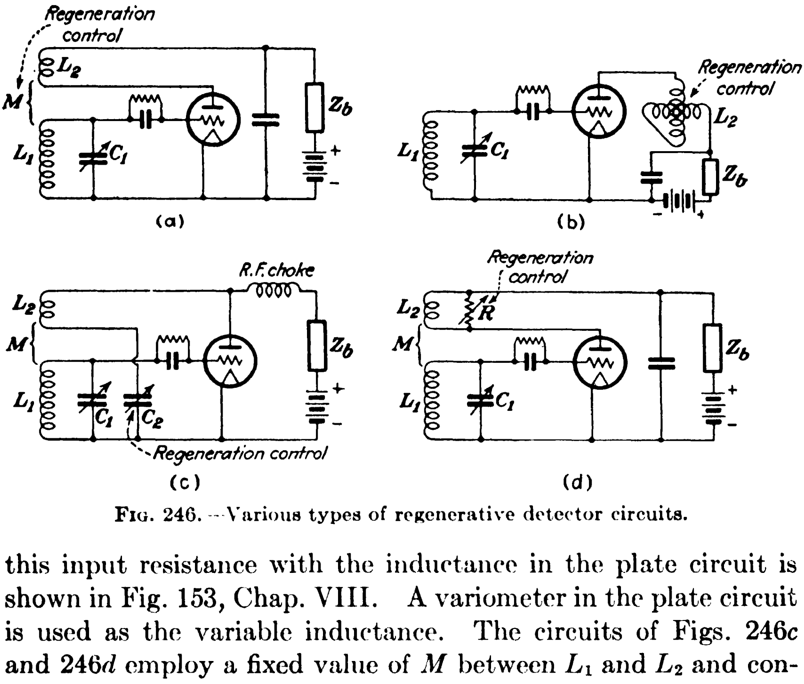


FIG. 246.—Various types of regenerative detector circuits.

this input resistance with the inductance in the plate circuit is shown in Fig. 153, Chap. VIII. A variometer in the plate circuit is used as the variable inductance. The circuits of Figs. 246*c* and 246*d* employ a fixed value of M between L_1 and L_2 and control regeneration in the one case by means of a variable condenser C_2 in series with the tickler coil, and in the other by means of a variable resistance shunted across it.

The feed-back adjustment exerts a small effect upon the tuning of the input circuit, since the input capacitance of the tube is a function of the plate-circuit impedance. A greater amount of feed-back is required for critical regeneration at the lower frequencies, which causes the circuit to break into oscillation whenever C_1 is tuned to a higher frequency without first reducing the amount of feed-back. The constants of the grid leak and

condenser in the earlier types of regenerative receiving sets were not suitable in many cases for the oscillating condition, and often caused the oscillations to be interrupted at an audio-frequency rate owing to the periodic blocking of the detector tube. This caused an unpleasant howl to be produced in the loud-speaker. Furthermore, whenever oscillations occurred, the set behaved as a miniature transmitting station as the detector was coupled directly to the antenna circuit in these receiving sets. These oscillations produced objectionable interference with other listeners and constituted a serious problem in thickly populated areas, particularly since the common practice was to cause the detector to oscillate when distant stations were sought. Greater sensitivity was obtained with an oscillating detector and the heterodyne action enabled the carrier waves of distant stations to be located by the characteristic beat-frequency whistles heard in the loud-speaker as the tuning dial was rotated.

Regeneration can readily increase the effective value of Q of the associated tuned circuit to the point where serious frequency distortion is produced by the increased sharpness of resonance. This resultant discrimination against the side-band frequencies would be objectionable in broadcast reception. With short waves, the width of the side bands is a very much smaller percentage of the carrier frequency and the sharpening of the response curve by regeneration is correspondingly less serious. Distortion of this character is of no consequence in the reception of code signals, and regenerative detectors are commonly used in this field. High sensitivity can be more conveniently secured with radio-frequency amplification ahead of the detector and without the need of the critical adjustments required by regenerative detector circuits. Consequently, these circuits are no longer used to their former extent.

149. Superregeneration.[8]—If a regenerative detector is adjusted so that it is on the verge of oscillation a small impulse applied to the grid may cause an oscillation to begin. Once started, the oscillations will continue and will build up to an amplitude limited by the tube characteristics. Under this

[8] E. H. ARMSTRONG, Some Recent Developments of Regenerative Circuits, *Proc. I.R.E.*, vol. 10, p. 244, August, 1922; also, H. ATAKA, Superregeneration of an Ultra-short-wave Receiver, *Proc. I.R.E.*, vol. 23, p. 841, August, 1935.

condition the circuit resistance is negative and will continue to remain so, even though the initiating impulse or signal has subsided. If the circuit resistance can be made to vary from negative to positive values at a sufficiently rapid rate in such a manner as to maintain a small average positive value, this instability could be avoided and the very high regenerative sensitivity obtainable at the threshold of oscillation could be usefully employed. This is accomplished in superregenerative detectors either by inserting a low value of radio frequency of about 20 kc in the plate circuit, or by applying a similar potential to the grid. This quenching frequency may be supplied either by a separate oscillator or by the superregenerative tube itself. Typical cir-

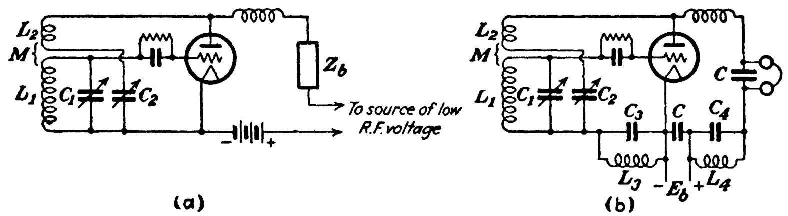


FIG. 247.—Typical superregenerative circuits.

cuits are shown in Fig. 247. In (*b*) the quenching frequency is produced by L_3C_3 and L_4C_4, while in (*a*) a separate oscillator is used.

The greater the difference between the quenching and carrier frequencies, the greater the sensitivity, as the signal has time to build up to a greater value during the half-cycle of the quenching frequency in which the circuit resistance is negative. The adjustments are rather critical and the selectivity is poorer than in the ordinary regenerative circuit, as the superregenerative circuit does not discriminate against signals of different frequency. The chief application of superregeneration is in the reception of signals at very short wave lengths where tuned radio-frequency amplification is at present impractical. In this region the lack of good selectivity is not so important as would be the case in the congested channels of lower frequency.

150. Vacuum-tube Voltmeters.—The rectifying properties of the vacuum tube enable it to be used to measure alternating voltages extending from the lowest audio frequencies up to radio frequencies of several megacycles. In its simplest form shown in

Fig. 248*a* the tube operates as a plate-circuit rectifier of the square-law type. The grid is biased negatively to almost cut-off and a suitable direct-current milliammeter is connected in the plate circuit. When an alternating potential E is applied to the grid, the average value of plate current as read by the direct-current instrument increases from I_b to I'_b. The device may be calibrated at 60 cycles and can be relied upon to be fairly accurate at the higher radio frequencies if the wave form is similar to that used in calibration. The input resistance is extremely high, and so the amount of energy abstracted from the source that is being measured is practically negligible. The input capacitance will be $C_{gf} + C_{gp}$, as the plate by-pass condenser is virtually a short circuit between plate and filament for alternating currents.

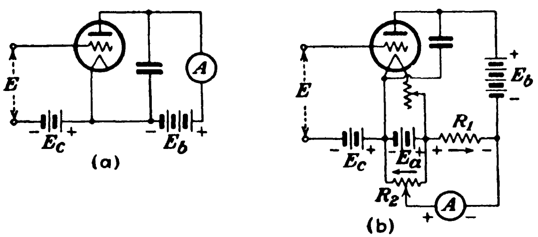


FIG. 248.—Vacuum-tube voltmeter circuits.

The base of the tube is sometimes removed in order to reduce the additional capacitance and dielectric loss contributed by this item.

The chief objection to the circuit of Fig. 248*a* is that the meter is given an initial deflection caused by the normal plate current I_b when no alternating voltage is applied to the grid. The arrangement of Fig. 248*b* avoids this difficulty by balancing out the initial value of plate current and enables a more sensitive direct-current instrument to be used. A resistance R_1 of about 2000 ohms is inserted in the plate circuit which will have a voltage drop of I_bR_1 across it of the polarity shown. The negative terminal of the direct-current microammeter is connected to the end of R_1 and the other instrument terminal is connected to the arm of a potentiometer R_2 shunted across the filament-heating battery. The potentiometer arm is then adjusted for no deflection of the meter. When an alternating voltage is impressed on the grid of the tube, the increase in plate current

through R_1 increases the voltage drop across this resistance and causes the meter to deflect.

For moderate values of impressed voltage the meter deflections are almost exactly proportional to the square of E. When this is the case, we may write

$$E = K\sqrt{D} \tag{34}$$

where K is some constant and D is the meter deflection. Then instead of determining the impressed voltage from a calibration curve, the constant K may be determined and the voltages corresponding to various deflections can be quickly computed on a slide rule. This will be found to be more convenient and rapid than using a calibration curve, particularly if a large number of readings are involved. The value of K can be determined by plotting E against $\sqrt{D}$ and determining the slope of the resultant straight line. Deviations from a linear relationship at the higher values of E will indicate the range for which (34) is valid.

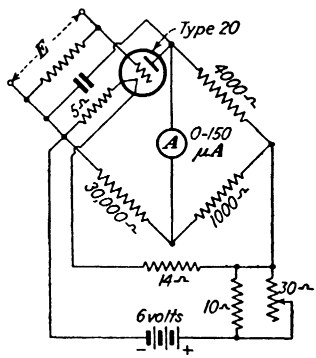


FIG. 249.—Bridge-type vacuum-tube voltmeter requiring only a single battery.

The circuits of Fig. 248 assume a conduction path through the voltage being measured so that the bias E_c will be actually applied to the grid. If not, a high resistance of 5 or 10 megohms shunted across the input terminals of the voltmeter will provide the necessary path.

Another form of vacuum-tube voltmeter which requires only a single battery is shown in Fig. 249.[9] The direct-current plate resistance R_p of the tube constitutes the fourth arm of the bridge which is initially balanced by adjusting the rheostat in series with the battery when no alternating voltage is applied. When an alternating voltage E is impressed, the value of R_p changes, unbalancing the bridge and causing the meter to deflect. A full-scale reading on a 150-scale microammeter is obtained with about 5 volts impressed on the grid.

[9] S. C. HOARE, A New Thermionic Instrument, *Trans. A.I.E.E.*, vol. 46, p. 541, 1927.

Where larger voltages are to be read, the reflexed arrangement of Fig. 250 may be employed. A biasing resistance R_c is used which increases the negative bias applied to the tube as E becomes larger. This is caused by the increase in plate current flowing through R_c, which is common to both grid and plate circuits. By using a large enough value of biasing resistance a full-scale reading of several hundred volts may be secured. The resulting calibration curve of the instrument becomes more nearly linear as the value of R_c is increased.

The calibrations of all of the above types of vacuum-tube voltmeters are more or less affected by the wave shape of the applied voltage. This is particularly true when even harmonics are present which cause the positive and negative loops of the wave to be dissimilar. Under this condition different deflections will be obtained for direct and reversed connections of E. In this respect these devices differ from thermocouple instruments, which always produce a deflection which is proportional to the r.m.s. value of the quantity being measured, regardless of its wave shape. This disadvantage is offset by the negligible amount of power required by vacuum-tube voltmeters, as in many types of measurements at audio and radio frequencies the power available is not sufficient to actuate a thermocouple instrument.

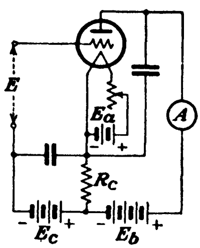


FIG. 250.—Reflexed vacuum-tube voltmeter.

Problems

1. A type 27 triode is operated as a plate-curvature detector with a pure resistance load of 10,000 ohms in the plate circuit. With 90 volts impressed on the plate and a negative grid bias of −9 volts on the grid, the d-c plate current is 0.63 ma and the tube constants are: $r_p = 21{,}000$ ohms, $r'_p = -550$, $\mu = 9$. What is the audio-frequency output voltage for a radio-frequency signal applied to the grid which has a maximum amplitude of 0.1 volt when unmodulated and a modulation factor of 0.5?

2. What will be the audio-frequency output voltage in Problem 1 if the resistance in the plate circuit is shunted with a by-pass condenser having negligible reactance for the carrier frequency? Assume the by-passing effect of the condenser to be negligible at audio frequencies.

3. The above tube is operated as a grid-leak detector with $E_b = 45$ volts and has an audio-frequency transformer in the plate circuit having a primary inductance of 100 henrys. The other conditions of operation are as follows: $\mu = 9$, $r_p = 9000$ ohms, $r_g = 150{,}000$ ohms, $v = 0.25$, $C_c = 250$ $\mu\mu$f,

$R_c = 1$ megohm, $C_g' = 15\ \mu\mu f$, $C_g = 50\ \mu\mu f$. The signal voltage across the tuned input to the detector is the same as in Problem 1. What will be the audio-frequency voltage across the transformer for the following modulating frequencies: (*a*) 100; (*b*) 1000; (*c*) 5000 cycles? Assume the effect of the by-pass condenser across the transformer primary to be negligible at audio frequencies.

4. In Problem 3, what will be the values of the second-harmonic voltages across the transformer? What are their relative percentages in terms of the fundamental voltages?

5. Repeat Problems 3 and 4, assuming the capacitance of the by-pass condenser across the primary of the audio-frequency transformer is 250 $\mu\mu f$.

E	I_bR_b (volts)	E	I_bR_b (volts)
0	20.0	9	90.5
1	21.7	10	101.4
2	26.0	11	112.3
3	32.5	12	123.6
4	40.6	13	135.0
5	49.9	14	146.0
6	59.5	16	168.6
7	68.5	18	191.5
8	80.0	20	214.0

6. The experimental data for a type 56 triode operated as a grid-bias detector are given in the accompanying table, where E is the maximum value of the applied carrier voltage. The test circuit used was the same as Fig. 232, with $E_b = 250$ volts, $E_c = -20$ volts, and $R_b = 0.5$ megohm. Assuming the maximum value of the unmodulated carrier to be 10 volts, determine the percentage of second harmonic in the audio-frequency output across R_b for the following modulation factors: $m = 0.25$, 0.5, 0.75, 1.0. Assume that the reactance of the by-pass condenser C is high enough to be neglected at audio frequencies.

7. Repeat Problem 6 for an unmodulated carrier of 4 volts. Compare these results with those of Problem 6 and explain why the amount of distortion is different.

8. A type 199 triode connected as in Fig. 248*a* is used as a vacuum-tube voltmeter, with $E_b = 45$ volts and $E_c = -4.5$ volts, $I_b = 240\ \mu a$, $\mu = 6.3$, $r_p = 36{,}700$ ohms, $r_p' = -2270$. Assuming the resistance of the meter in the plate circuit to be negligible in comparison to r_p of the tube, what will be the meter reading if an alternating potential of 1 volt, maximum value, is impressed on the grid?

CHAPTER XIII

RECEIVING SYSTEMS

151. Types of Receivers.—The two most commonly used systems of radio reception are the tuned radio-frequency receiver and the superheterodyne. The regenerative circuit, consisting of a regenerative detector followed by an audio-frequency amplifier, is hardly ever used now, although at the outset of broadcasting this was the standard type of circuit. The inadequate selectivity of only one tuned coupled circuit preceding the detector has rendered it obsolete for broadcast reception. Regenerative detectors are still used to a considerable extent in short-wave receivers where their ability to oscillate is an advantage in the reception of telegraphic code signals. However, the detector is usually preceded by at least one stage of tuned radio-frequency amplification in order to increase the selectivity and sensitivity, and to prevent radiation.

The superregenerative circuit, described in the preceding chapter, is of use in the field of very short waves. These wave lengths are being used for police communications from transmitter-equipped patrol cars and in other applications where the limited range of these very high frequencies is of no disadvantage.

152. Tuned Radio-frequency Receivers.—This type of receiving set became widely used as soon as satisfactory neutralizing methods were developed to prevent self-oscillations in the radio-frequency amplifier. The first sets had individual tuning dials for each radio-frequency transformer, which made the tuning adjustments somewhat inconvenient. These were followed shortly by designs using a gang control of the tuning condensers, enabling a single tuning dial to be used. The introduction of screen-grid tubes a few years later did away with the need of neutralizing circuits. The same number of tuned stages had to be used with these tubes in order to obtain the necessary selectivity, which resulted in a greatly increased amount of gain in the amplifier. About this same time grid-leak detection in broadcast

receivers began to be replaced by power detection, mostly of the grid-bias type, which reduced the distortion inherent in square-law detectors. The lower sensitivity of this method of detection was more than offset by the additional gain of the screen-grid tubes.

With the advent of screen-grid pentodes for radio-frequency amplification, still higher gains per stage became possible. These tubes have an amplification factor μ of from 1200 to 1300 and a mutual conductance of about 1600 micromhos. This value of g_m is of greater significance than μ, as the resonant amplification of a tuned transformer-coupled amplifier of the type illustrated in Fig. 251 has been shown in Chap. IX to be given by

$$A_v = \frac{E_2}{e_g} = g_m Q_2 \omega M \qquad (1)$$

where $Q_2 = \omega L_2/R_2$, and r_p of the tube is very large compared to ωM. As a result of the larger amplifications possible, pentodes of this type have practically superseded screen-grid tetrodes as radio-frequency amplifiers in receiving sets.

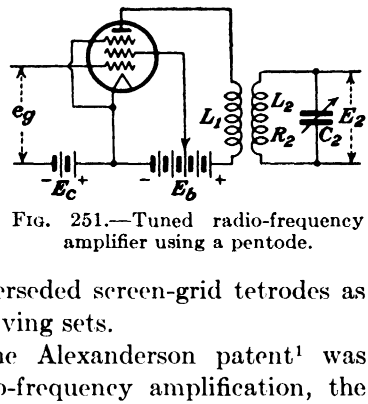


FIG. 251.—Tuned radio-frequency amplifier using a pentode.

As the expiration date of the Alexanderson patent[1] was approached covering tuned radio-frequency amplification, the owners of this patent began to license others to manufacture superheterodyne receivers, the patents of which they also controlled. This has brought about a marked decrease in the production of tuned radio-frequency sets in favor of the more sensitive and selective superheterodyne.

153. The Superheterodyne.—In the superheterodyne or double detection circuit the incoming frequency is superimposed on a slightly higher or lower frequency generated by a local oscillator and the resultant combination is then impressed upon a detector The rectifying action of the latter extracts the beat or intermediate frequency which is then amplified by the intermediate frequency amplifier. This intermediate frequency will contain all of the modulation frequencies present in the original signal, the action of the heterodyne detector having in effect converted the original carrier frequency into a new carrier frequency much

[1] U. S. patent 1,173,079.

lower in value. The output of the intermediate-frequency amplifier is then impressed on a second detector, which is followed by an audio-frequency amplifier. The schematic arrangement of the circuit elements is shown on Fig. 252. The second oscillator associated with the second detector is required only when continuous-wave code signals are to be received. The dots and dashes are heterodyned by adjusting this oscillator so that its frequency differs from the intermediate frequency by some convenient amount, such as 1000 cycles. The 1000-cycle beats produced are then rectified and made audible by the second detector. The output of the second oscillator can be introduced anywhere in the intermediate-frequency amplifier. This oscil-

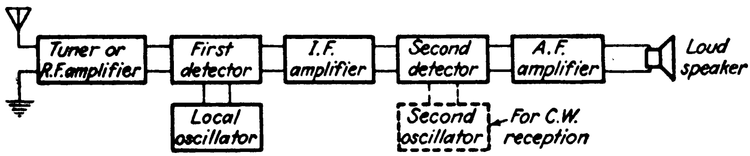


FIG. 252.—Schematic arrangement of the elements in a superheterodyne receiver.

lator, while not necessary for radio-telephone reception, is sometimes used in multirange broadcast receivers to assist in picking up and tuning in short-wave signals. When used in this manner, the oscillator frequency is made the same as the intermediate frequency of the set and produces a high-pitched audible note in the loud-speaker when the fringe of the signal is first reached. The pitch becomes lower as resonance is approached, enabling the set to be adjusted to exact resonance by tuning the beat frequency to zero. The second oscillator is then switched off.

The superheterodyne was originally designed to overcome the difficulties encountered when the amplification of frequencies higher than a few hundred kilocycles was attempted. Satisfactory methods of securing stability had not as yet been developed and amplifiers operating at these frequencies would oscillate at the slightest provocation. By converting the frequency of the received signal into one of much lower value, fairly high gain could be obtained without instability. In the earlier superheterodynes the intermediate frequency used was low enough to enable resistance-coupled amplifiers to be employed.[2] Resist-

[2] E. H. ARMSTRONG, A New System of Short Wave Amplification, *Proc. I.R.E.*, vol. 9, p. 3, 1921.

ance coupling was soon replaced by suitable iron-core transformers which greatly increased the amount of gain. Modern practice usually employs tuned air-core transformers, although tuned transformers with a special core of molded iron dust are being used to some extent.

By tuning one or more stages of the amplifier to the chosen value of intermediate frequency the superheterodyne becomes extremely selective. Suppose the intermediate-frequency amplifier is tuned to 30 kc and a desired signal of 1000 kc is to be received in the presence of an interfering signal of 1010 kc, or a value 1 per cent higher. Adjusting the local oscillator to 1030 kc will produce the required beat frequency of 30 kc, while that portion of the interfering signal which manages to get through the tuned circuits ahead of the first detector will produce a beat frequency of 20 kc, or 33⅓ per cent lower than the resonant frequency of the intermediate-frequency amplifier. As a result of the heterodyning process, two signals differing in frequency by only 1 per cent will now differ from each other by a matter of 33⅓ per cent, making it very much easier to discriminate against the unwanted signal than would have been the case with the original frequency separation of only 1 per cent. The lower the value of intermediate frequency used, the greater the heterodyne selectivity becomes.

However, the use of intermediate frequencies as low as that of the above example, while formerly employed, is too low for modern broadcast reception. This is due to troubles caused by image-frequency interference and the lack of sufficient frequency stability on the part of the local oscillator when single-dial control of both oscillator and tuned input circuits to the first detector is attempted. It also becomes difficult to avoid undue attenuation of the outer-side-band frequencies in the intermediate-frequency amplifier unless the stages of the latter are coupled by means of suitable band-pass filters. If 30 kc is used as an intermediate frequency for broadcast reception, the response curve of the amplifier should be reasonably flat from 25 to 35 kc and then fall off sharply on either side of this band. This characteristic, while attainable, requires the accurate adjustment of an appreciable number of circuit elements which would be costly from a production standpoint. Once adjusted, however, the constants would not have to be altered as the desired signals

are all converted to the same value of intermediate frequency by the local oscillator. By using a higher intermediate frequency the width of the band to be uniformly transmitted becomes a much smaller percentage of the resonant frequency, and the desired characteristics can be obtained from simple tuned coupled circuits having characteristics similar to Fig. 61, page 87. These considerations as to fidelity do not apply to the reception of telegraphic-code signals.

With the advent of tubes that are entirely stable without the need of neutralizing circuits, much higher values of intermediate frequency can be used and image-frequency interference can be more readily avoided.

154. Image-frequency Interference.—This term is given to the interference produced in superheterodyne receivers caused by stations which differ in frequency by twice the value of the intermediate frequency. In the previous example an interfering signal of 1060 kc would have also produced 30-kc beats with the local oscillator frequency of 1030 kc, and hence would have been amplified and impressed upon the second detector. If the oscillator frequency is changed slightly, one of the two beat frequencies present will rise while the other will fall, and the beats between them will be rectified by the second detector and will be heard as a characteristic heterodyne whistle. Since the oscillator frequency is usually higher than the received-signal frequency in broadcast receivers, a signal in order to produce image-frequency interference must have a frequency of $f + 2F$, where f is the frequency of the desired station and F is the value of the intermediate frequency used.

The remedy for this form of interference is to provide adequate selectivity ahead of the first detector, usually in the form of one or more stages of tuned radio-frequency amplification. This was seldom done in the older types of superheterodynes, with the result that as broadcasting stations increased in number and power, it became almost impossible to tune in a station without having a heterodyne whistle in the background owing to image-frequency interference. Even with several tuned circuits preceding the first detector, objectionable image-frequency interference may be caused by strong local stations at various settings of the tuning dial. A method of balancing out this interference is shown in Fig. 253. The antenna circuit contains an inductance

L_a of sufficient size to lower the resonant frequency of the circuit below the lowest frequency to be received. The antenna is coupled to the first tuned circuit L_1C_1 by means of an adjustable condenser C_a. Magnetically coupled to L_1 is a second tuned circuit L_2C_2, the proper magnitude and sign of M being obtained by two reversed turns on L_1. A very small portion of the signal producing image interference is impressed upon L_2C_2 from the antenna by means of an adjustable condenser C of a few micro-microfarads. By proper choice of the sign of M the interfering voltage introduced in this manner by C can be made to be in phase opposition to that which manages to pass through the tuned circuits. The set is then tuned to a location where image-

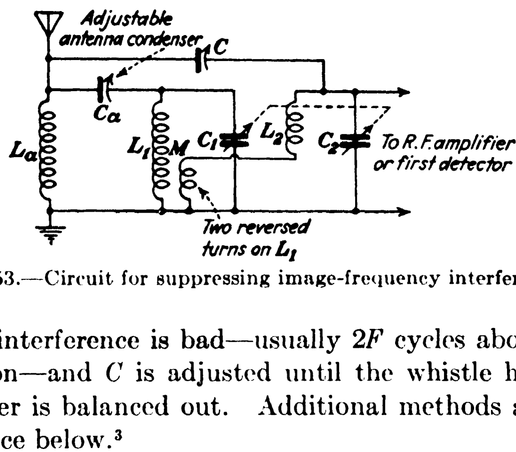


FIG. 253.—Circuit for suppressing image-frequency interference.

frequency interference is bad—usually $2F$ cycles above a strong local station—and C is adjusted until the whistle heard in the loud-speaker is balanced out. Additional methods are given in the reference below.[3]

The use of a higher value of intermediate frequency will reduce image-frequency interference, as the burden imposed upon the selective circuits preceding the first detector becomes correspondingly less. These tuned circuits must be adequately shielded, or local signals capable of producing image-frequency interference may be picked up directly by the wiring, resulting in a greatly reduced image-frequency ratio. This is particularly true of the input circuit to the first detector.

155. Choice of Intermediate Frequency.—Low values of intermediate frequency, while giving good adjacent-channel selectivity, are troublesome because of image-frequency interference.

[3] H. A. WHEELER, Image Suppression in Superheterodyne Receivers, *Proc. I.R.E.*, vol. 23, p. 569, June, 1935.

A widely used value for broadcast receivers has been 175 kc. More recently, with the advent of multirange receivers for short-wave reception, higher frequencies such as 465 kc have become fairly common. This value, while high so far as broadcast frequencies are concerned, is inconveniently low when reception at 20 megacycles is considered. Here, oscillator and signal frequencies differ by only 2.3 per cent, which is hardly enough for a good image-frequency ratio. Also, since oscillator and signal-frequency circuits are all tuned by a common control on short waves, the frequency stability of the oscillator must be fairly good. A further difficulty is experienced in that an appreciable voltage of oscillator frequency may appear in the signal-frequency circuits owing to stray couplings when these two frequencies differ from each other by only a small percentage. A frequency in the neighborhood of 465 kc seems to be the best compromise for multirange sets, as any value much higher will either fall within or be too close to one of the tuning ranges.

Another factor to be guarded against in the choice of an intermediate frequency is to avoid those which are being actively used in other branches of radio communication, as there is a possibility of these signals being picked up by unshielded portions of the intermediate-frequency amplifier and first detector.

156. Frequency Converters.—The production of the intermediate frequency formerly required separate tubes for the oscillator and first detector or "mixer," as it is sometimes called. A typical arrangement is shown in Fig. 254. The oscillator is of the tuned-grid type, using the voltage applied to the screen-grid of the detector tube as a source of plate supply. The plate coil L_p is usually wound on top of the grid coil L_g to insure tight coupling between them. The tube is connected across only a portion of L_g so as to minimize the effect of tube variations on the oscillator frequency. Self-bias is obtained by means of a 40,000-ohm grid leak and a 250-$\mu\mu$f condenser.

The oscillator voltage may be induced directly into the tuned input of the detector by providing magnetic coupling between L_1 and L_g, or else by inducing the required voltage in series with the cathode circuit of the detector by means of coil L_2, as shown in the diagram. The latter method reduces the reaction between the oscillator and signal-frequency circuits and is preferable for this reason. The detector is self-biased by means of a 7000-ohm

resistor in the cathode circuit. This negative bias may be further increased by means of the volume control which supplies an adjustable biasing potential between the grids and cathodes of the other radio-frequency and intermediate-frequency amplifiers, as a means of controlling the overall sensitivity of the set. Increasing the negative bias of the detector will reduce its output of intermediate-frequency, enabling it to assist in the control of sensitivity. This requires the use of a "variable-mu" tube having a remote cut-off, the details of which will be described later.

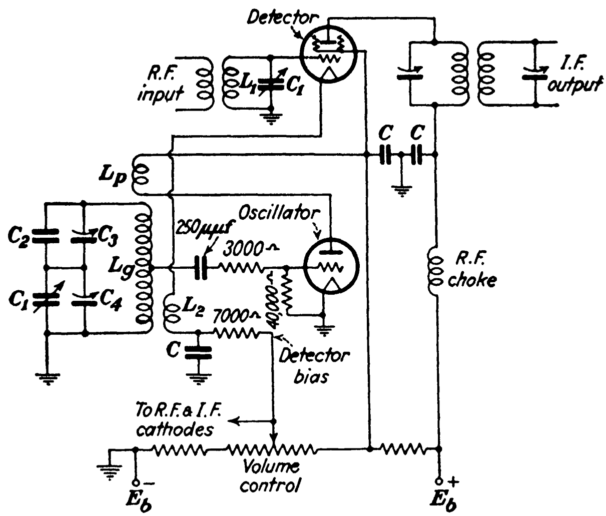


FIG. 254.—Frequency converter for a superheterodyne consisting of separate oscillator and detector tubes.

In order for a constant frequency-difference to exist between the oscillator and signal-frequency circuits when a single tuning control is used, either the oscillator condenser plates must have a special shape, or, if both condensers are of equal capacitance, a combination of series and shunt condensers must be used. The latter arrangement is shown in Fig. 254. The oscillator tuning condenser C_1, which is mounted on the same shaft and identical with that across L_1, is in series with a fixed capacitance made up of C_2 and C_3 in parallel. The trimmer condenser C_3 is for the purpose of accurately adjusting the combination to the desired value. It can be shown that with this scheme the oscillator will

"track" with practically a constant frequency-difference from the resonant frequency f_1 of L_1C_1. The relation desired is

$$F = f - f_1 \tag{2}$$

where F is the intermediate frequency and f is the oscillator frequency. This relationship is exactly obtained at only three frequencies within the tuning range. These three points are called "tracking frequencies" and can be arbitrarily selected by the designer. One of these is usually chosen near the center of the frequency band to be covered and the other two near, but not at the ends of the band. By their proper selection the departure of F from its desired constant value will not be more than a few kilocycles at points between and beyond these tracking frequencies. The flat-topped response curves of the intermediate-frequency transformers will permit a nominal variation in the intermediate frequency F to be tolerated over the tuning range. With the series capacitance $C_2 + C_3$ omitted, exact tracking will occur at only two points instead of three.

Typical values of the required tuning constants of the circuit of Fig. 254 for an intermediate frequency of 175 kc are as follows:

Tuning range	550 to 1500 kc
Tuning condenser C_1	15 to 350 $\mu\mu$f
R-F transformer L_1	270 μh
Oscillator tuning inductance L_o	210 μh
Series capacitance $C_2 + C_3$	836 $\mu\mu$f
Trimmer capacitance C_4	3.78 $\mu\mu$f
Tracking frequencies	600 kc 1000 kc 1400 kc

157. Electron-coupled Frequency Converters.[4]—Instead of coupling the oscillator to the first detector or mixer by some reactive means, the electron stream within the mixer can be modulated by the oscillator frequency. This greatly reduces the possibilities of reaction between oscillator and signal-frequency circuits. Tubes designed for this purpose, known as "pentagrid converters," have both oscillator and mixer elements all contained within the same bulb, which does away with the need of a separate oscillator tube.

[4] C. F. Nesslage, E. W. Herold, and W. A. Harris, A New Tube for Use in Superheterodyne Frequency Conversion Systems, *Proc. I.R.E.*, vol. 24, p. 207, February. 1936.

The electrode arrangement is shown in Fig. 255. Grid 1 and grid 2 serve respectively as the control grid and anode of the oscillator. These two electrodes, in conjunction with the cathode, may be viewed as composite cathode furnishing an electron supply which has been modulated at oscillator frequency. Grids 3 and 5, which are connected together within the tube, function in much the same manner as a screen grid, except that they electrostatically shield the signal-frequency control grid (grid 4) instead of enclosing the plate as in the ordinary screen-grid tube. Since this screen is at a positive potential with respect to the cathode, it aids in accelerating the modulated electron stream toward the plate. The signal-frequency voltage applied to grid 4 further modulates the electron stream so that the plate current will contain the various sum-and-difference combinations of the oscillator and signal frequencies. The tuned intermediate-frequency transformer then transmits the desired band of frequencies and excludes the others. The difference frequency between oscillator and signal frequency is the one nearly always

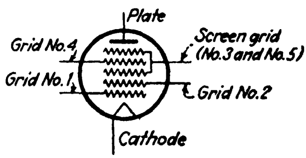


FIG. 255.—Electrode arrangement in a pentagrid converter.

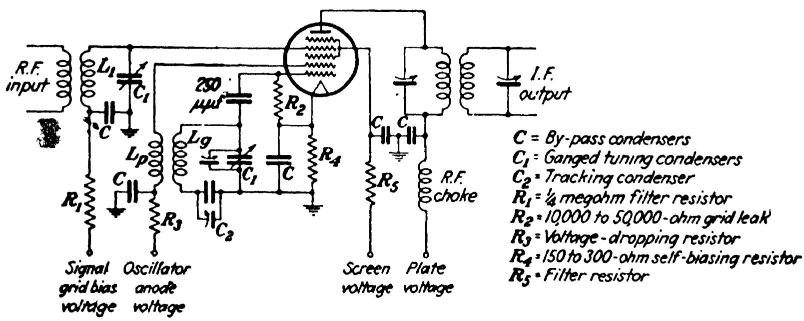


FIG. 256.—Typical pentagrid converter circuit.

used, although the sum of these two could be used instead, if desired.

A typical pentagrid converter circuit is shown in Fig. 256. The oscillator circuit is similar to that of Fig. 254, the tracking being obtained in the same manner. The sensitivity of the mixer portion of the tube is controlled by means of a variable negative bias applied through R_1 to grid 4, usually obtained from

some form of automatic volume control. This bias is in addition to that of about 3 volts produced by the total cathode current flowing through R_4. The mixer has the remote cut-off characteristic of variable-mu tubes for convenience of sensitivity control.

Pentagrid converters are capable of producing a translation gain of from 50 to 100, this gain being defined as the intermediate-frequency output voltage divided by the radio-frequency input voltage. In the same way as the mutual conductance (or transconductance) of a tube may be used as an indication of its merit as an amplifier, the *conversion transconductance* s_c of a mixer tube may be defined as the ratio of an increment of the intermediate-frequency current in the primary of the *i-f* transformer, to a corresponding increment of the radio-frequency signal voltage applied to the control grid. This term can be used in frequency-converter computations in exactly the same way as g_m in the case of single-frequency amplifiers. The value of s_c diminishes as the negative bias applied to the control grid of the mixer tube is increased.

158. Variable-mu Tubes.—The usual method of controlling the gain of a radio-frequency amplifier is to impress an adjustable negative bias on the tube grids. However, when strong local signals are received, the bias must be adjusted almost to cut-off in order to reduce the mutual conductance sufficiently to avoid overloading the tube in the last stage. Not only does this make the setting of the gain or volume control extremely critical, but "cross talk" and distortion of the modulated envelope are also apt to be serious when the operating potential of the grid is shifted into this region of high curvature. One form of cross talk due to curvature of the characteristic has already been mentioned in Sec. 139. This particular form, it will be recalled, is similar to the action of a superheterodyne; one of the two stations involved serves as the local oscillator, while the other may be regarded as producing the signal. The initial amplifier tube functions as the first detector, and both stations may be heard whenever the amplifier is tuned to either the sum or difference of the frequencies of the two stations. The tuned amplifier and detector of the receiving set serve as intermediate-frequency amplifier and second detector, respectively.

A second form of cross talk may also occur when a strong interfering signal is impressed upon an amplifier that is tuned to

a desired signal of different frequency. This form of interference is due to the modulation of the desired carrier by the modulation frequency of the interfering signal, and occurs whether the desired carrier is modulated or not. The manner in which this takes place can be best understood by a mathematical formulation of the conditions. Assume the alternating component of the amplifier plate current to be represented by the series

$$i_p = c_1 e_g + c_2 e_g^2 + c_3 e_g^3 + \cdots \tag{3}$$

The constants in (3) can be evaluated by the method of (24), Chap. VI, giving

$$i_p = \frac{di_p}{de_g} e_g + \frac{1}{2} \cdot \frac{d^2 i_p}{de_g^2} e_g^2 + \frac{1}{6} \cdot \frac{d^3 i_p}{de_g^3} e_g^3 + \cdots + \frac{1}{\lfloor n} \cdot \frac{d^n i_p}{de_g^n} e_g^n \tag{4}$$

Let

$$e_1 = E_1 \sin \omega_1 t \tag{5}$$

be the voltage impressed on the grid due to the desired signal, assumed to be unmodulated for simplicity, and let the modulated interfering signal be

$$e_2 = E_2(1 + m_2 \sin \omega_s t) \sin \omega_2 t \tag{6}$$

The total alternating voltage acting on the grid will be

$$e_g = e_1 + e_2 \tag{7}$$

Substituting this value of e_g in (4) will enable the various frequency terms in the output to be determined. Terms having frequencies in the vicinity of ω_1 will be produced by the odd powers of e_g in (4). These will be

$$(i_p)_{\omega_1} = \frac{di_p}{de_g} e_1 + \frac{1}{6} \cdot \frac{d^3 i_p}{de_g^3}(e_1^3 + 3e_2^2 e_1) + \frac{1}{120} \cdot \frac{d^5 i_p}{de_g^5}(e_1^5 + 10 e_2^2 e_1^3 + 5 e_2^4 e_1) + \cdots \tag{8}$$

The first term is evidently the desired signal obtained by ordinary amplifier action and would be the only one present if the characteristic were linear. Substituting the values of e_1 and e_2 in the second term of (8) gives

$$(i_p)_{\omega_1} = \cdots + \frac{d^3 i_p}{de_g^3}\left[\frac{1}{6}E_1^3 \sin^3 \omega_1 t + \frac{1}{2}E_2^2 E_1(1 + m_2 \sin \omega_s t)^2 \sin^2 \omega_2 t \cdot \sin \omega_1 t\right] + \cdots \quad (9)$$

Expanding the last term in the brackets,

$$(i_p)_{\omega_1} = \cdots + \frac{1}{2}\cdot\frac{d^3 i_p}{de_g^3}E_2^2 E_1(1 + 2m_2 \sin \omega_s t + m_2^2 \sin^2 \omega_s t)(\tfrac{1}{2} - \tfrac{1}{2}\cos 2\omega_2 t) \sin \omega_1 t + \cdots \quad (10)$$

one term of which is

$$(i_p)_{\omega_1} = \cdots + \frac{1}{4}\cdot\frac{d^3 i_p}{de_g^3}E_2^2 E_1(1 + 2m_2 \sin \omega_s t) \sin \omega_1 t + \cdots \quad (11)$$

The term given by (11) will be recognized to be of the same form as that of a standard modulated signal, and evidently the signal-frequency term ω_s of the interfering station is modulating the carrier of the desired station. The fifth and higher derivatives will contribute similar terms. This second form of cross talk, like that of the first type, may be reduced by providing adequate selectivity ahead of the first amplifying tube. The use of untuned coupling circuits between the first tube and the antenna should not be used if cross talk is to be avoided.

Tetrodes are more troublesome in this respect than triodes, particularly on strong signals, since, with the customary method of gain control, the higher amplification of these tubes requires that the operating point be shifted closer to cut-off in order to prevent overloading of the last amplifier tube. In this region of high curvature the higher order derivatives are much larger than at normal values of bias. This formerly required the use of a manually operated "local-distance" switch which reduced the sensitivity of the amplifier for local reception by means other than increased negative bias on the control grid. Since automatic volume control varies the grid bias to control sensitivity, objectionable cross talk and distortion are apt to be experienced with strong signals. This led to the development of the variable-mu or "supercontrol" type of tube having a remote cut-off.[5]

[5] STUART BALLANTINE and H. A. SNOW, Reduction of Distortion and Cross-talk in Radio Receivers by Means of Variable-mu Tetrodes, *Proc. I.R.E.*, vol. 18, p. 2102, December, 1930.

The modification in the characteristic which this construction brings about may be thought of as having been produced by two tubes in parallel, the resultant characteristic being given by $A + B$ in Fig. 257. Tube A has a high value of μ while tube B has a low value. When weak signals are to be received, the operating bias would be at point x where the mutual conductance is high, while with strong signals the operating point would be shifted to y where g_m is much lower. The distortion produced by operating at a point on curve A having the same mutual conductance—*i.e.*, the same slope—as at point y would evidently produce appreciable distortion if the signal voltage were large. Instead of using two tubes, the resultant characteristic of Fig. 257 can be obtained with a single tube by winding the control grid so that a portion of it has a coarser mesh than the remainder of the grid. As the potential of the grid is made increasingly

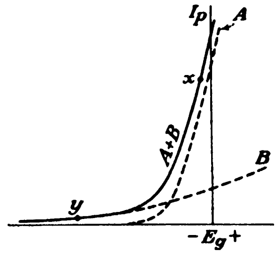


FIG. 257.—Composite characteristic obtained by operating two tubes in parallel having different values of amplification factor.

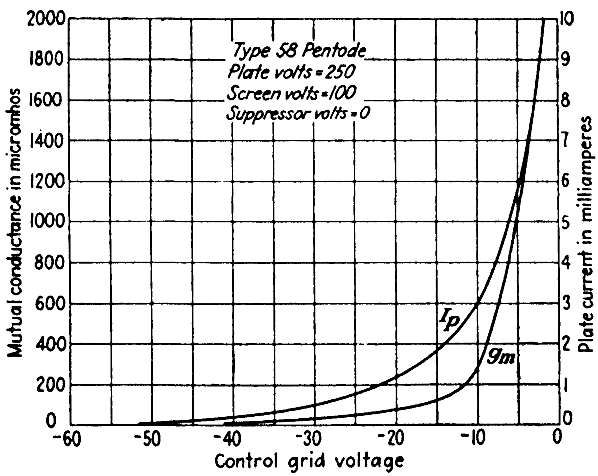


FIG. 258.—Characteristics of a typical variable-mu pentode.

negative, the electron current passing through the finer-mesh portion of the grid is rapidly reduced, and finally becomes zero. A reduced number of electrons continue to flow through the coarser portion of the grid. Various other modifications in the

electrode structure could also be used to secure this type of characteristic.[5]

Most of the tetrodes and pentodes designed for reception purposes in radio-frequency amplifiers are of the variable-mu type. The characteristics of a typical variable-mu pentode are shown in Fig. 258. The mutual conductance falls rapidly as the negative bias on the grid is increased, but beyond -10 volts the decrease in g_m is quite gradual. This type of characteristic enables automatic volume control to be successfully applied, since these tubes do not need to be biased to practically cut-off in order to reduce their sensitivity sufficiently for local reception, with the resultant reduction of distortion and cross talk.

159. Automatic Volume Control.—As previously mentioned, this system of volume control automatically varies the mutual conductance of the various amplifying tubes (other than audio frequency)—usually by imposing a variable negative biasing voltage between the control grid and cathode—in such a manner as to vary the sensitivity of the tubes so as to maintain essentially constant carrier voltage across the detector. The output from the loud-speaker will then vary only with the degree of modulation of the received signal and not with variations in the field strength of the received carrier caused by fading. Strictly speaking, this should be called "automatic sensitivity control," as the term "volume control" is preferably applied to means of controlling the gain of the audio-frequency amplifier.

Automatic volume control (abbreviated a.v.c.) is accomplished by utilizing the direct-current output voltage of a detector as an additional control-grid bias on the various radio-frequency amplifier tubes. In superheterodynes the mixer tube is also included. This rectified control voltage may be obtained from either the detector tube itself, or from an auxiliary rectifier tube. Diodes are usually employed for this purpose, and in duplex tubes having two diode elements, one of them may be used as the detector (the second detector in superheterodynes), while the other diode element can be used as the a.v.c. tube.

The manner in which this automatic control of sensitivity is brought about may be seen from Fig. 259. The radio-frequency voltage appearing across the tuned circuit LC is rectified by the

[5] *Loc. cit.*

diode, the rectified current producing an IR drop of the polarity shown across the resistor R_1. Assuming the cathodes of the preceding radio-frequency amplifying tubes to be at ground potential, a negative bias can be impressed upon the grids of these tubes by connecting their grid-return leads to the negative end of R_1 through the high resistance R_2. The latter is for the purpose of preventing undesired coupling between the amplified

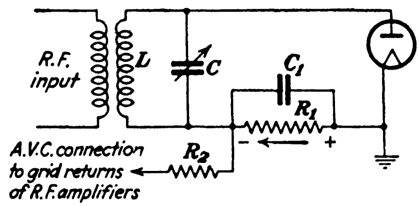


FIG. 259.—Diode rectifier used to secure automatic volume control.

output and the grid circuits of the preceding amplifier tubes, and has a value in the vicinity of a megohm. If the strength of the received signal increases, the voltage drop across R_1 rises, biasing the amplifier grids more negatively and thus reducing their gain. Conversely, a weaker signal reduces the amount of negative bias, causing the amplifier gain to increase. In this way the sensitivity of the receiver is altered so as to hold the loud-speaker

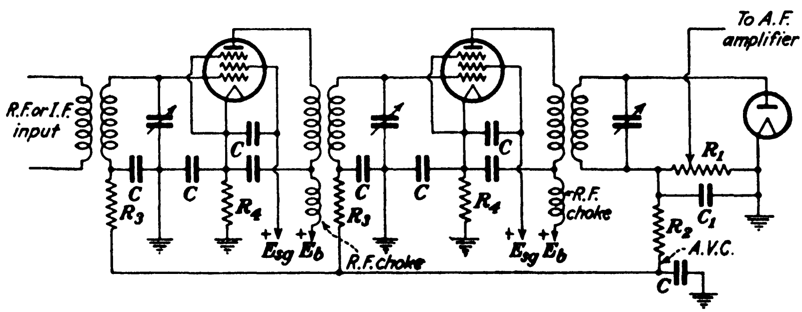


FIG. 260.—Automatic volume control applied to radio-frequency amplifier.

output at normal volume. When no signal is being received, there would be zero bias on the amplifier tubes thus controlled, which would be objectionable. This may be avoided by the use of individual biasing resistors R_4 in the cathodes of the various tubes to be controlled which provide a fixed minimum bias, as shown in Fig. 260. The same tube is used for detection as well

as automatic volume control in this case. The filtering resistors R_3, in conjunction with the by-pass condensers C, serve to prevent feed-back couplings between stages.

The circuit just described will begin to function as soon as any signal whatever is received. It is usually desirable to delay the controlling action until the received signal exceeds a definite minimum value so as to retain the maximum sensitivity of the amplifier for the reception of very weak signals. This can be accomplished by impressing a negative bias E_c on the diode plate, as shown in Fig. 261. While a battery is shown as the source of this biasing voltage, the IR drop across a suitable bleeder resistance in the B supply would actually be used. No current will flow in the diode circuit unless the amplitude of the signal voltage across LC exceeds the value of E_c. A separate detector unit must now be used as no rectifying action

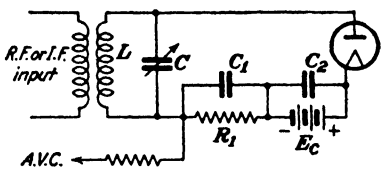


FIG. 261.—Circuit for delayed automatic volume control.

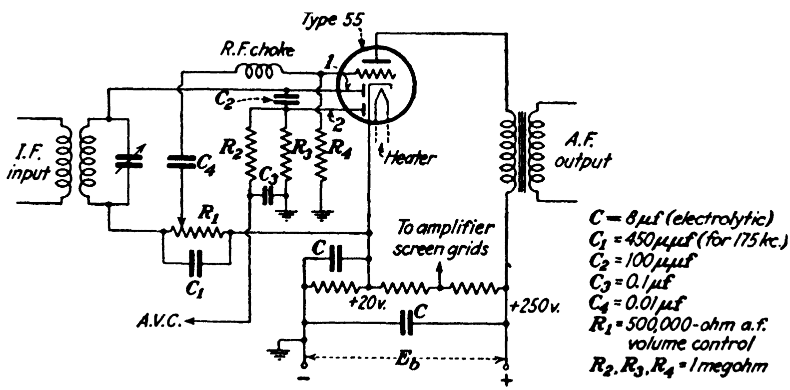


FIG. 262.—Circuit employing duplex-diode triode tube for detector, delayed automatic volume control, and audio amplifier.

will take place in the a.v.c. tube with weak signals incapable of developing a voltage across LC greater than E_c.

A suitable circuit for this purpose is illustrated in Fig. 262. A duplex-diode triode tube is used for the combined functions of second detector, delayed a.v.c., and audio-frequency amplifier. Half-wave-detector action is obtained by plate 1 of the first diode

unit. A portion of the rectified voltage developed across R_1, which serves as a volume control, is impressed on the grid of the triode unit through C_4 in series with the radio-frequency choke coil. The latter aids in excluding the intermediate frequency from the audio-frequency amplifier. A 20-volt fixed negative bias is impressed on the triode by connecting its grid to ground through a 1-megohm resistor R_4, as the common cathode is held positive with respect to ground by this amount by means of the voltage drop across a portion of the bleeder resistance. This same bias also serves as the voltage E_c in Fig. 261 for delayed automatic volume control.

Automatic volume control is obtained by means of the second diode element which is connected to the first unit by means of C_2. The rectified output of this second unit flows through R_3, the drop across which is applied to the grid-return leads of the preceding amplifiers through the additional filtering elements R_2C_3. The two diode elements are essentially in parallel across the secondary of the intermediate-frequency transformer.

Diode biasing of the triode unit, as was done in Fig. 241, is impractical in Fig. 262 as the plate load of the triode now consists of the comparatively low transformer primary resistance, and moving the arm of the volume control to the right to reduce the volume would also reduce the C bias impressed on the tube, which would result in excessive values of plate current. Accordingly, a fixed value of bias must be used. It is interesting to note that reducing the volume by means of the adjustable arm on R_1 does not reduce the voltage developed across the intermediate-frequency transformer. Consequently, the detector unit of the diode is at all times supplied with the proper voltage for best operation; provided, of course, that the strength of the received signal is sufficient to overcome the delay feature of the automatic volume control.

The advantages of automatic over manual control of sensitivity are that variations in volume caused by fading of the received signals are practically eliminated, and that local stations may be tuned through without producing excessive blasts of sound, which are not only annoying but are apt to damage the loud-speaker. All stations which are of sufficient intensity to actuate the a.v.c. will produce equal volumes of sound from the loud-speaker regardless of signal strength, provided the signals are all modu-

lated to the same degree. The volume obtained will be governed by the setting of the volume-control arm on R_1. Automatic volume control is very useful in mobile receiving sets, such as auto radios, where reception conditions may be changing continuously owing to the shielding effects of buildings and overhead wiring.

One of the disadvantages of the system is that tuning is made more difficult as exact resonance can no longer be determined by adjusting the tuning dial for maximum volume. The volume will remain constant for an appreciable movement of the tuning dial so that the dial must be adjusted for the best quality or the minimum background of noise. Visual aids in the form of suitable meters or neon glow tubes, which are actuated by the a.v.c. biasing voltage, are sometimes incorporated in the set. More recently, a special form of cathode-ray tube has been used for this purpose. Another drawback is that the sensitivity of the set rises to a maximum when no signal is being received so that the user receives the full benefit of the prevailing noise and static while passing between the various carrier waves in search of a desirable program. Circuits designed to suppress this interchannel noise are discussed in the following section.

160. Interchannel-noise Suppression.—Circuits for the automatic suppression of noise when no signal voltage is impressed on the a.v.c. tube use a separate control tube which cuts out the audio-frequency amplifier when the signal falls below a certain minimum value. The principle usually employed is illustrated by the schematic circuit of Fig. 263.

Tube A is a diode rectifier, tube B is the noise-suppression-control or "squelch" tube, and C is the first audio-amplifier tube. When a signal voltage is present across LC, a rectified current flows through R_1 which fluctuates with the degree of modulation. A portion of the resultant voltage drop across R_1 is applied to the grid of tube B through the filtering elements R_2C_2. This voltage drop is sufficient to bias tube B beyond cut-off—E_1 being small—so that the plate current through R_3 is reduced to zero. Tube C will then have normal bias impressed on its grid from E_2 and will amplify the fluctuating voltages existing across R_1, a portion of which are applied to its grid through C_4 which is connected to the arm of the volume control. But when no signal is being received, there is no rectified voltage

across R_1, and hence no bias is imposed on tube *B*. The plate current of this tube then rises to its maximum value which produces an *IR* drop across R_3 of the polarity shown, which, in conjunction with E_2, is sufficient to bias tube *C* beyond cut-off and prevents it from amplifying. The audio-frequency amplifier is thus cut out whenever the signal voltage falls to too low a value.

This feature can be combined with automatic volume control and diode detection by the use of a single tube of the duplex-diode triode type. One diode element is used as the detector while the other furnishes a.v.c. Noise suppression is accomplished by the triode element, which is connected across a portion of the load resistance in the detector circuit in the same manner

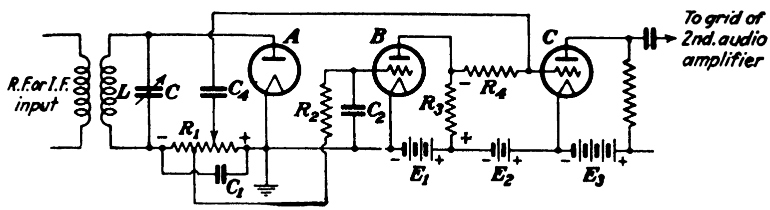


FIG. 263.—Automatic noise-suppression circuit.

as in Fig. 263. While it is possible to couple both diodes to the same source, it is usually more desirable to supply them with signal inputs obtained from separate sources having different degrees of selectivity, if the noise control is to act promptly when the set is detuned from the received signal by a moderate amount. If the a.v.c. diode is supplied from the more broadly tuned source, the gain of the controlled amplifier will not rise so rapidly with a moderate amount of detuning. At the same time the more sharply tuned supply to the detector is reduced more rapidly by the detuning, enabling the noise-control tube to function before the volume of noise has reached objectionable proportions. This broadly tuned source of signal input to the a.v.c. diode in superheterodynes can be obtained either from a tertiary winding tightly coupled to the primary of the input transformer feeding the second detector, or else from a broadly tuned auxiliary amplifier fed from the first detector. The latter method requires two intermediate-frequency amplifiers—one for the signal and one for the automatic volume control. This separate supply to the a.v.c. tube, while more complicated and expensive, is advantage-

ous even when noise suppression is not used, as the selectivity and sensitivity of the auxiliary amplifier can be designed to meet the particular requirements of the a.v.c. system. In general, the selectivity should be less, and the sensitivity greater, than that of the signal amplifier.

Automatic suppression of noise has the objection that when weak signals are being received, the audio-frequency amplifier may be cut in and out continuously as the signals fluctuate in intensity. For this reason a switch should be provided to prevent the noise-control tube from operating on weak-signal reception.

161. The Compandor.[6]—This device is a form of automatic volume control applied to the voice frequencies before modulation takes place at the transmitting station and after detection at the receiver, and is an important adjunct in the transoceanic radio telephone. It automatically compresses the range of the speech-energy variations at the transmitter and expands the range to normal at the receiving end, thus greatly improving the speech to noise ratio. The transmitting device is called the "compressor," the receiving device, the "expandor," and the complete system, the "compandor."

The total range of speech intensities impressed upon the transmitter from the connecting telephone line, if no form of volume control whatever were used, would vary perhaps as much as 70 db, or an energy ratio of 10 million to 1. Using manual gain control on the speech amplifier supplying the radio transmitter, so as to keep the radio output at a high average level at all times, would still show a considerable variation in signal intensity from syllable to syllable from the inherent nature of speech sounds. For example, the energy levels of some consonants in speech are 30 db below that of the stronger vowels. Consequently, if the speech input is adjusted so as to avoid overloading the transmitter on the high speech intensities, the weaker sounds may frequently fall below the prevailing atmospheric-noise level at the receiving end and be completely masked by static.

The function of the compressor is to reduce this range of speech-energy variations. This is accomplished in the manner shown in

[6] R. C. MATHES and S. B. WRIGHT, The Compandor—An Aid against Static in Radio Telephony, *Bell System Tech. Jour.*, vol. 13, p. 315, July, 1934, or *Elec. Eng.*, vol. 53, p. 860, June, 1934.

Fig. 264. The resistances R_1 and R_2 may be regarded as the series arms of an attenuating network or "pad," the shunt arm of which is made up of the two plate resistances in series of the push-pull tubes. The tubes are biased to nearly cut-off by E_c. A portion of the output is rectified, producing a voltage E_g of the polarity shown which is proportional to the envelope of the voltage E_2. The C bias on the two tubes is thus made to fluctuate in accordance with the output envelope, which in turn causes similar variations in the tube plate resistances. The resistance of the shunt arm of the attenuator is thus varied by an amount depending upon E_2; the greater E_2, the lower the value of the shunting resistance and the greater the amount of attenuation offered by

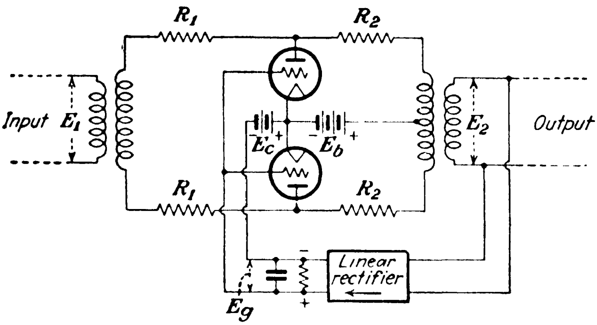


FIG. 264.—Compressor circuit of compandor.

the network. This distorted-speech output is then further garbled by suitable means, so as to insure the privacy of the conversation; following which, it becomes the speech input to the transmitter. The schematic details of the latter have already been discussed in connection with Fig. 220, page 335, for the long-wave channel of the transatlantic radio telephone.

At the receiving end the compressed waves are unscrambled by the receiving privacy device and sent into the expandor which inverts the operation of the compressor. The expandor circuit is shown in Fig. 265 and is merely a push-pull amplifier biased almost to cut-off by E_c. The total grid bias, which is $E_c - E_g$, is again made to vary in accordance with the rectified envelope of E_3. If the resistances R are kept small compared with the plate resistances of the control tubes, it may be shown that the output of this device will be just the inverse of the compressor operation and the speech will be restored to its original form.

The schematic arrangement of parts for the compandor and its associated equipment are shown in Fig. 266. The "vodas," taken from the initials of the words "Voice-Operated Device Anti-Singing," consists of voice-operated relays with their associated vacuum-tube detectors and delay networks. They are operated by the voice currents in the circuit so as to eliminate the

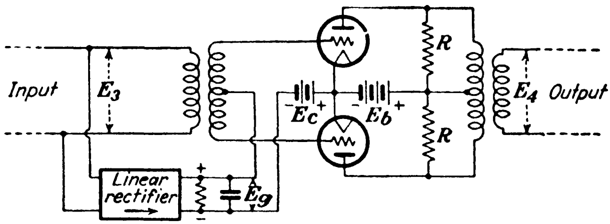


FIG. 265.—Expandor circuit of compandor.

possibility of singing at all times by blocking the feed-back paths. Their operating details are described in the reference below.[7]

The compandor has enabled the transatlantic radio telephone to be commercially usable through static-noise levels approximately 5 db higher than was formerly possible. This increases

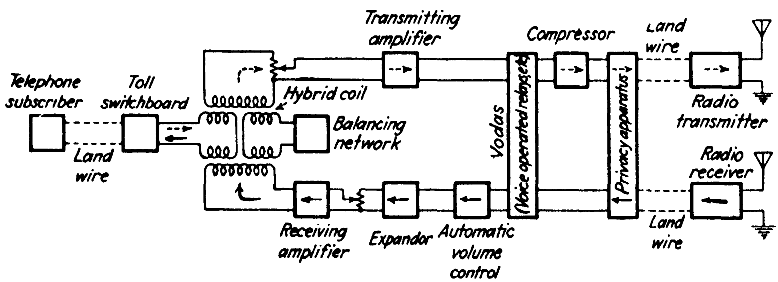


FIG. 266.—Compandor applied to one end of transoceanic radio-telephone.

the service time of the circuit, particularly during the seasons of the year when it is most needed.

162. Acoustically Compensated Volume Control.—The volume controls previously described vary the magnitude of the signal voltage applied to the audio-frequency amplifier without changing the frequency-response curve of the set. If the volume is progressively reduced, it is found that both the low and the high fre-

[7] S. B. WRIGHT and D. MITCHEL, Two-way Radio Telephone Circuits, *Proc. I.R.E.*, vol. 20, p. 1117, July, 1932; or *Bell System Tech. Jour.*, vol. 11, p. 368, July, 1932.

quencies tend to disappear, and at low volume levels only the middle register is heard. This effect is due to the greater sensitivity of the ear to sounds in the vicinity of 1000 cycles. A great many people prefer to have the volume-level of music from a radio set reduced to a point considerably lower than would be the case if the artists were actually present in the same room with the listeners. When the ordinary type of volume control is used, this practice results in an appreciable loss in the musical quality, as most of the bass notes will be practically inaudible.

In order to compensate for this effect, it is necessary to attenuate the middle range of frequencies much more rapidly than the higher and lower frequencies as the volume is reduced. A common method of achieving this is to shunt a broadly tuned series-resonant circuit across the lower portion of the volume-control resistance, as shown in Fig. 267. As the volume is reduced by moving the arm of the volume control downward, the shunting effect of the resonant circuit becomes more pronounced. The frequencies above and below resonance are affected to a much lesser extent. The resonant frequency chosen is usually in the vicinity of 1000 cycles, typical values for L and C being 0.05 henry and 0.5 microfarad. About 20 to 25 per cent of the total volume-control resistance is included between the terminals of L and C.

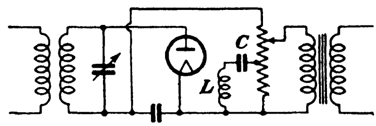


FIG. 267.—Acoustically compensated volume control.

163. Tone Control.—Most broadcast-receiving sets are equipped with a tone control to enable the users to alter the over-all-fidelity characteristic to suit their individual tastes. The ordinary form of tone control used consists of an adjustable resistance in series with a condenser, the combination being shunted across some portion of the audio-frequency amplifier, as shown in Fig. 268. The action is to increase the shunting effect of C_1 as R_1 is reduced and thus increase the attenuation of the higher frequencies.

When pentodes are used as power tubes, the output also is usually shunted by a condenser and resistance in series, to offset the rise in the load impedance at the higher audio frequencies caused by the inductance of the loud-speaker. In the cheaper sets the resistance R_2 is sometimes made variable to serve as the

tone control, and R_1C_1 is dispensed with. This is objectionable, as the effective load impedance will change with the tone-control adjustments, with the possibilities of causing appreciable amplitude distortion. Pentodes are much more critical in the matter of load impedance than triodes, as discussed in Sec. 81.

The principal advantage of a tone control that progressively removes the higher frequencies is in the reception of weak signals. When the signal strength is not much above the noise level due to static, removing the higher frequencies renders the noise less objectionable. Sporadic noise has a relatively wide frequency distribution and if the upper audio frequencies are suppressed, the fidelity of the music will be somewhat impaired, but the noise will be reduced in volume and rendered considerably less irritating in quality.

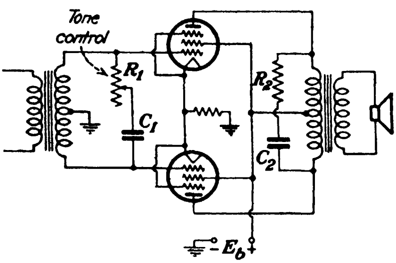


FIG. 268.—Tone-control circuit across amplifier input.

164. Typical Broadcast-receiving Sets.—The majority of modern broadcast receivers are of the superheterodyne type which are capable of more uniform selectivity and better performance than the tuned radio-frequency type. Superheterodynes are also more suitable for multichannel reception since the same intermediate-frequency amplifier can be used on all channels without circuit changes.

The circuit diagram of a typical broadcast receiver of the superheterodyne type is shown in Fig. 269. A single pentode stage of tuned radio-frequency amplification precedes the combined mixer and oscillator tube. Two tuned coupled circuits similar to Fig. 253 are sometimes used ahead of the first radio-frequency amplifier tube, particularly when a low value of intermediate frequency is employed, to secure adequate selectivity as a protection against image-frequency interference. A pentagrid converter is used as a first detector and oscillator. The circuit details are practically the same as those discussed in Sec. 157 in connection with Fig. 256. In all-wave sets a separate oscillator tube electron-coupled to a pentagrid converter is sometimes used. This aids in reducing possible reactions between oscillator and signal-frequency circuits on short-wave reception as there is a

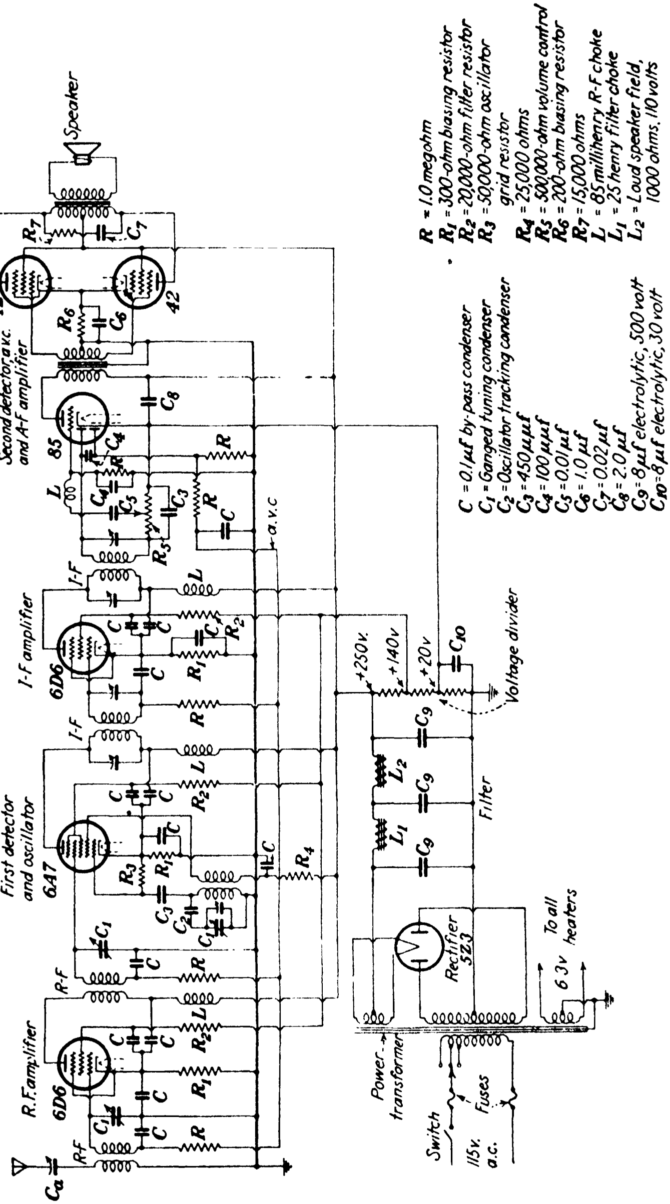


FIG. 269.—Circuit diagram of a typical broadcast receiver of the superheterodyne type.

tendency for volume-control variations of the control-grid bias to affect the amplitude of the oscillations in pentagrid converters because of the transconductance between control-grid and oscillator anode.[8] A shift in oscillator frequency may accompany this amplitude variation.

A single stage of intermediate-frequency amplification is used following the mixer. In the more expensive sets two such stages are sometimes employed, and although the full amount of gain thus afforded can be used only at times of extremely low atmospheric-noise level, increased selectivity is obtained by virtue of the additional tuned circuits.

The intermediate-frequency output is impressed on a duplex-diode triode, one plate of which serves as the second detector while the other furnishes delayed automatic volume control, as described in connection with Fig. 262. The a.v.c. is applied to the grids of the first three tubes through the resistors R. The triode portion of this same tube is the first audio-frequency amplifier and operates with a fixed bias of 20 volts obtained by connecting its grid to ground through a 1-megohm resistance. The common cathode is held 20 volts positive with respect to ground, as shown, in order to obtain the delayed a.v.c. action. The triode output is transformer-coupled to a pair of push-pull pentodes, which are self-biased by resistor R_6. The types of output tubes and their mode of operation will differ with various sets; push-pull Class A, or Class AB, being the most common. In some cases two or more loud-speakers are used, which results in better acoustic performance.

Since all tubes with the exception of the rectifier use indirectly heated cathodes of the same voltage rating, all of the heaters may be connected in parallel and operated from one secondary winding on the power transformer. The two ends of the high-voltage secondary are connected to the two plates of a full-wave rectifier, the rectified output of which is impressed across the filter. The series inductances L_1 and L_2 of this structure offer a high impedance to the flow of the alternating current while the shunt capacitances offer a low-impedance path, the effect becoming progressively greater for the harmonics of the rectified output.

[8] W. A. HARRIS, The Application of Superheterodyne Frequency Conversion Systems to Multirange Receivers, *Proc. I.R.E.*, vol. 23, p. 279, April, 1935.

The two filter sections shown are sufficient to reduce the residual alternating voltage across the voltage divider to negligible proportions. The larger the values of inductance and capacitance employed, the greater the filtering action. Electrolytic condensers are practically always used. One of the two choke coils is usually the field winding of the dynamic loud-speaker.

Multirange or all-wave receivers alter the inductances of the radio-frequency and oscillator coils by suitable switching arrangements so as to use the same gang-tuning condensers for all channels. This may be accomplished either by using tapped coils, or preferably, by switching in individual coils or transformers for each range. The unused lower-frequency coils are short-circuited so as to prevent the possibility of absorption effects due to the natural periods of these coils when not shunted by the gang-tuning condenser. Each of the transformer secondaries is provided with individual trimmer condensers—and tracking condensers in the case of the oscillator coils—which enables them to be properly aligned.

165. Performance Tests on Receiving Sets.—The properties of sensitivity, selectivity, and fidelity of a receiver may be determined by means of tests which have become fairly well standardized.[9] These tests consist of impressing a known modulated signal on a standard artificial antenna connected to the receiver under test and measuring the audio-frequency output into a pure resistance load which replaces the loud-speaker. Standard output is taken as 50 milliwatts at 400 cycles into a resistance load adjusted to the value recommended by the tube manufacturer to give maximum undistorted output for this type of tube under the circuit conditions used in the set. The ratio of transformation of the output transformer must be taken into consideration in arriving at the value of load resistance to be used.

The sensitivity at a particular frequency is defined as the effective value of a radio-frequency signal voltage modulated 30 per cent at 400 cycles that will produce standard output when inserted in the artificial antenna. The constants of the standard antenna for broadcast frequencies are $R = 25$ ohms, $L = 20\ \mu h$, $C = 200\ \mu\mu f$, and the effective height is assumed to be 4 meters. The circuit arrangement is shown in Fig. 270. The standard

[9] I.R.E. Year Book, p. 121, 1931.

signal generator consists of a modulated oscillator with a calibrated attenuator. The voltage input to the attenuator is read by a suitable thermocouple meter and the output voltage is determined by the setting of the attenuator. These output voltages will ordinarily range from perhaps a volt to values less than a microvolt. This range requires very careful shielding and filtering of the signal-generator circuits in order to prevent stray pick-up by direct radiation when a sensitive receiver is tested.

The selectivity can be determined by varying the carrier frequency impressed upon the set and determining the increase in voltage necessary to produce standard output, the modulating frequency and percentage of modulation remaining the same as for the sensitivity test. Selectivity characteristics for a typical broadcast receiver of the superheterodyne type are shown in Fig. 271. It will be noted that the selectivity is fairly uniform throughout the tuning range, the variation being caused by the tuned radio-frequency amplifier ahead of the first detector. This variation in selectivity is very much less than with tuned radio-frequency receivers.

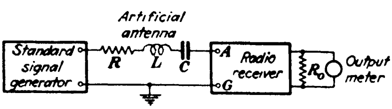


FIG. 270.—Circuit arrangement for determining receiver performance.

The fidelity can be determined by adjusting the receiver and signal generator for standard output at 400 cycles at a particular carrier frequency and then varying the modulating frequency applied to the signal generator. Fidelity curves for the receiver used in Fig. 271 are shown in Fig. 272. The difference between the curves for the two different carrier frequencies is caused by the increased attenuation of the outer side bands due to the slightly greater selectivity at 600 kc. The reduction in response beginning at about 1000 cycles is due chiefly to this factor, augmented by the radio-frequency by-pass condenser across the detector load resistance. The drop in response at the lower audio frequencies is caused by the audio-frequency amplifier.

Overall-fidelity tests are sometimes made by measuring the sound output by means of a suitable calibrated microphone. The amplified output of the microphone is then measured by means of either a vacuum-tube voltmeter or a thermocouple instrument. This form of test has the advantage of including the loud-speaker characteristics. The chief difficulty with tests

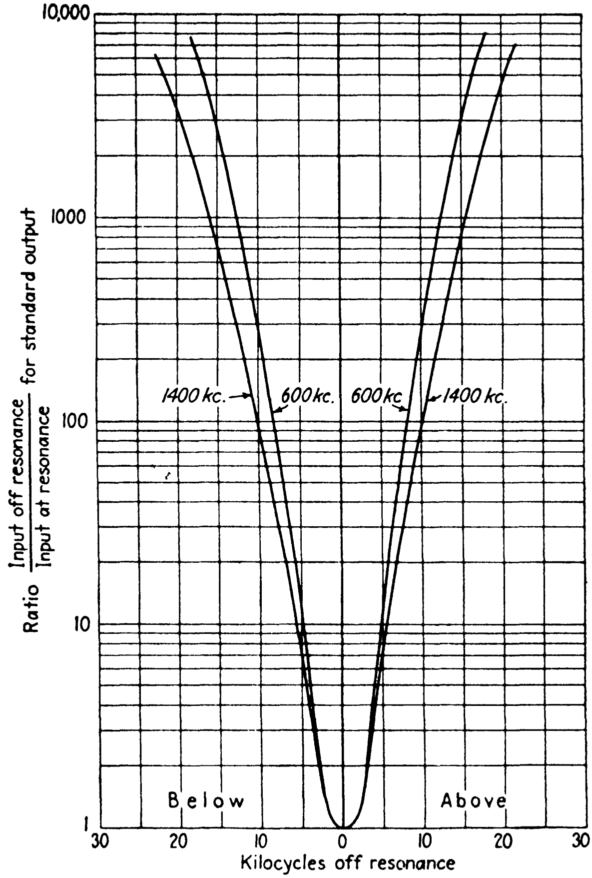


FIG. 271.—Selectivity curves for a typical superheterodyne receiver.

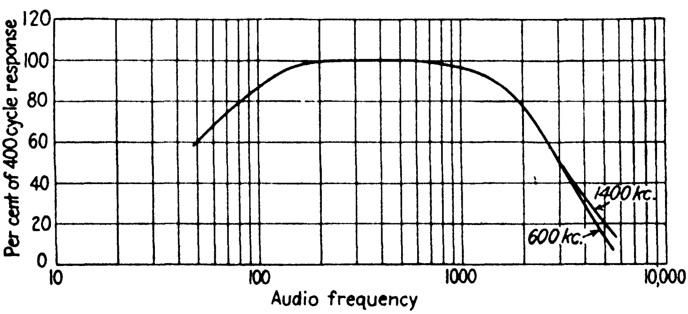


FIG. 272.—Fidelity curves for a typical superheterodyne receiver.

of this nature is that the acoustics of the room in which the test is conducted will modify the results to a very considerable degree. Reflections from the walls, ceiling, etc., will produce standing sound waves in the room, the pattern of which changes continuously with the frequency, causing the microphone output to vary in an erratic fashion as the frequency is varied. These difficulties may be partly overcome by the liberal application of sound-absorbing materials to the various reflecting surfaces. The use of a directional microphone of the ribbon type is also of assistance in reducing the effects of sound reflection. If the microphone is made to oscillate back and forth by suitable mechanical means, it can be made to average the sound-wave pattern and give an indication of what the sound intensity would be if reflections were absent. This averaging can be conveniently accomplished by using a rather sluggish thermocouple meter which is unable to follow the fluctuations caused by the movement of the microphone. At low frequencies the distance between the nodes and antinodes of the sound-wave pattern becomes larger so that this method of averaging the sound output becomes impractical. The loud-speaker may be tested separately in this same manner by applying audio frequencies to the input terminals of the device.

166. Loud-speakers.—The usual type of loud-speaker consists of a paper cone driven at the apex by a suitable acoustic motor. The horn type of loud-speaker is seldom used for broadcast reception because of its poor fidelity when restricted to the small sizes that can be accommodated in the ordinary radio cabinet. The larger horns have excellent fidelity and when actuated by a suitable driving mechanism may have an acoustic efficiency in the vicinity of 30 per cent.[10] These types are commonly used in talking pictures.

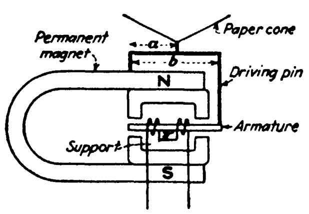


Fig. 273.—Balanced armature type of loud-speaker.

One type of actuating mechanism that has been widely used is illustrated in Fig. 273. A soft-iron armature is held between the pole pieces of a permanent magnet by a flexible support. Sur-

[10] E. C. Wente and A. L. Thuras, *Bell System Tech. Jour.*, vol. 7, p. 140, January, 1928.

rounding the armature with sufficient clearance so as not to interfere with its motion are two fixed coils which carry the audio-frequency current. This current causes the armature to vibrate and its motion is communicated by the driving pin to the cone through the linkage shown. This linkage is in effect a mechanical transformer whose transformation ratio is b/a, and serves to alter the mechanical impedance of the load as viewed from the driving source, in much the fashion as an output transformer in the electrical case, and for the same purpose.

The amplitude of vibration is limited in this device and in order to secure reasonable output at low frequencies, a large diameter of cone is required. This adversely affects the performance at high frequencies as the cone will no longer vibrate as a piston but breaks up into complex modes of vibration. The inductance of the coil is relatively high, causing the impedance to rise from a few thousand ohms at low frequencies to perhaps as high as 40,000 ohms at 5000 cycles. This impedance variation causes the power output from the amplifier to vary appreciably with the frequency.

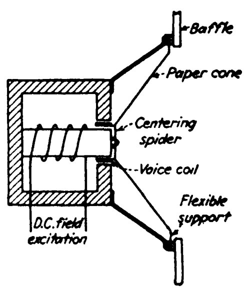


FIG. 274.—Moving coil or "dynamic" loud-speaker.

The moving coil or electrodynamic loud-speaker has largely replaced the magnetic type just described, and is shown in Fig. 274.[11] A paper cone is driven by a voice coil of a small number of turns located in a strong radial magnetic field. When traversed by an alternating current, the coil is subjected to an alternating force which is proportional to BlI, where B is the flux density, l is the length of wire on the voice coil, and I is the current. The cone, which may be from 5 to 12 in. in diameter, is centered by means of a suitable flexible spider. A similar flexible support is provided at the outer edges of the cone, these two supports furnishing the elastic restoring forces. The impedance of the voice coil is low, usually of the order of 5 to 15 ohms, and the phase angle is comparatively small. This results in a fairly constant impedance throughout the range of audio frequencies and insures

[11] This type was first described by C. W. Rice and E. W. Kellogg, *Trans. A.I.E.E.*, vol. 44, p. 982, 1925.

a more constant transfer of power from the amplifier over the frequency range.

The magnetic field is produced in most cases by using the field coil as one of the choke coils in the B supply filter, as mentioned in connection with Fig. 269. A permanent magnet may also be used where the energy for field excitation is not available, or is too expensive, as with battery-operated sets.

At frequencies below 1000 cycles the cone vibrates as a piston and in order to secure adequate sound output at low frequencies, it is necessary to employ a baffle so as to prevent the radiations from the front and back sides of the cone (which are 180 degrees out of phase) from canceling each other. The baffle should be large enough to make the distance from the front to the back of the cone greater than one-quarter wave length for the lowest frequency with which we are concerned. The cabinet serves as a baffle in broadcast receivers, but in the smaller types the dimensions are entirely too small. As a result, these sets are usually deficient in the bass register.

A comprehensive discussion of loud-speakers and their characteristics is beyond the scope of this chapter. For detailed information on this subject the reader is referred to "Applied Acoustics," by H. F. Olsen and F. Massa, Chap. VII.

Problems

1. Using the data given on page 388 for the superheterodyne frequency-converter circuit of Fig. 254, compute the deviations produced in the intermediate frequency from the desired value of 175 kc when the input circuit to the detector is tuned to 800, 1200, and 1500 kc. The oscillator and detector tuning condensers C_1 are mounted on the same shaft and have identical capacitances at all dial settings. Neglect tube-input capacitance and inductive couplings to L_1 and L_g.

2. It is desired to determine the constants of a frequency converter similar to Fig. 254 for one channel of an all-wave superheterodyne receiver to cover the receiving range from 10 to 23 megacycles, using an intermediate frequency of 465 kc. At 10 megacycles the maximum capacitance of the tuning condenser is to be used, which is 400 $\mu\mu$f. The oscillator frequency is to be higher than the signal frequency and the tracking frequencies are to be 12, 16, and 20 megacycles. Find the necessary values of L_1, L_g, $C_2 + C_3$, and C_4.

CHAPTER XIV

ANTENNAS AND WAVE PROPAGATION

167. Electromagnetic Waves.—The existence of electromagnetic waves was predicted by Maxwell in 1865. By mathematical reasoning he showed that variable currents in a conductor produce electromagnetic waves in space, that these waves travel with the velocity of light, and that light itself consists of electromagnetic waves of very short wave lengths. Maxwell's theory was verified in 1888 by Hertz, who was the first to produce experimentally and detect these waves. He was able to perform with them many of the standard experiments in optics and found that as the wave length was shortened, their properties became more and more analogous to the properties of light. The experiments of Hertz were duplicated by other investigators and as the apparatus and technique improved, it became possible to detect these waves at distances of several hundred feet. This suggested the idea of wireless telegraphy and a patent for such a system of communication was applied for by Marconi in 1896.

If a conductor is traversed by an alternating current, it will be surrounded by an alternating magnetic field. As this current builds up from zero, the strength of the magnetic field increases, and it will be in time phase with the current in the immediate vicinity of the conductor. At a point remote from the conductor the rise and fall of the field will lag behind the phase of the current because of the finite velocity of propagation of the field. This velocity of propagation is 3×10^8 meters—approximately 186,000 miles—per second, the same as for light. Therefore, if the frequency of the current in the conductor is a million cycles per second, the current will have risen from zero to a maximum and back to zero again during the time that the initial magnetic impulse has traveled 150 meters. This is depicted in Fig. 275 for an elementary length of wire at O carrying a current of $i = I_m \sin \omega t$. The magnetic vector H lying in a plane passing through O and perpendicular to the conductor will have traveled

from O outward to a, or one-half wave length, during this time. The vectors to the right of point a were produced by the preceding half cycle of current which was flowing in the opposite direction in the conductor. The field at any point p a distance r from the conductor will be given by

$$H = H_m \sin \omega\left(t - \frac{r}{c}\right) \tag{1}$$

where c is the velocity of light and H_m is the maximum value of the magnetizing force at the point p.

These vectors representing the instantaneous values of H at various points may be thought of as moving away from the con-

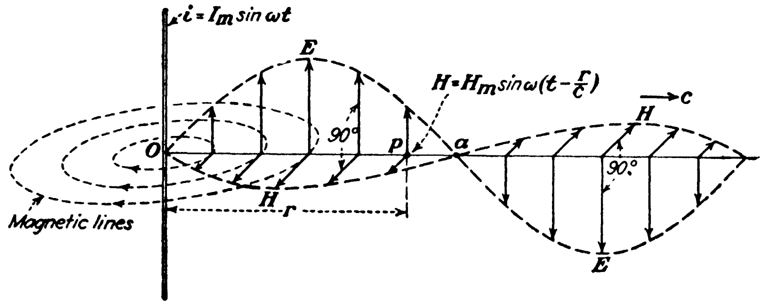


FIG. 275.—Production of electromagnetic waves in space.

ductor with the velocity of light and hence will be capable of inducing an e.m.f. in other conductors in the vicinity which may be cut by this magnetic field. However, a conductor is not necessary, as a vertical rod of insulating material would also be cut by these magnetic lines of force, and would have the same difference of potential produced between its two ends as a conducting rod of the same dimensions. Evidently a column of air would undergo the same experience. Consequently, a moving electromagnetic field will give rise to a difference of potential in space so that an electrostatic field will be produced. This is represented in Fig. 275 by the electric vectors E. These vectors are in *time phase* with H, but they will be in *space quadrature*, as shown.

A uniform magnetic field of B lines per square centimeter moving with a velocity v cm. per second will induce an e.m.f. in a wire l cm. long and at right angles to B of

$$E = Blv \text{ abvolts} \tag{2}$$

If B and l are both unity and $v = 3 \times 10^{10}$ cm. per second, the velocity of light, the electrostatic voltage gradient $\mathcal{E}$ will be 3×10^{10} abvolts per centimeter. In air a value of H gilberts per centimeter will produce a flux density of B lines per square centimeter, the permeability μ being unity, so that $B = H$; and when the latter is moving with the velocity of light, as in the case of electromagnetic waves, the relationship between $\mathcal{E}$ and H is

$$\mathcal{E} = 300H \tag{3}$$

168. Electrical Units.—Two systems of electrical units arise in the c.g.s. system, the *electrostatic* (e.s.u.), and the *electromagnetic* (e.m.u.). The former is developed from the fundamental definition that a unit charge is one which repels a like charge at a distance of 1 cm. with a force of 1 dyne. The electromagnetic system of units is derived from the definition of a unit magnetic pole as one which will repel a like pole at a distance of 1 cm. with a force of 1 dyne. From these fundamental definitions the various other electrical units may be derived. The relationship between the two systems of units involves the velocity of light c, or the value squared, as shown in Table I.

TABLE I

Practical unit		C.G.S. electromagnetic units (e.m.u.)		C.G.S. electrostatic units (e.s.u.)	Ratio e.m.u./e.s.u.
1 coulomb	=	10^{-1} abcoulomb	=	3×10^9 statcoulombs	c
1 ampere	=	10^{-1} abampere	=	3×10^9 statamperes	c
1 volt	=	10^8 abvolts	=	$\frac{1}{3} \times 10^{-2}$ statvolt	$1/c$
1 ohm	=	10^9 abohms	=	$\frac{1}{9} \times 10^{-11}$ statohm	$1/c^2$
1 farad	=	10^{-9} abfarad	=	9×10^{11} statfarads	c^2
1 henry	=	10^9 abhenrys	=	$\frac{1}{9} \times 10^{-11}$ stathenry	$1/c^2$

Both systems of units are employed in electromagnetic theory, electrical quantities being expressed in e.s.u. and magnetic quantities in e.m.u. Such a composite system of units is called the *Gaussian* system. In this system (3) becomes

$$\mathcal{E} = H \tag{4}$$

where $\mathcal{E}$ is in statvolts per centimeter and H is in gilberts per centimeter

169. Current and Voltage Distribution in an Antenna.—An antenna possesses distributed inductance and capacitance and is electrically similar in behavior to a long transmission line, such as a telephone line with the distant end on open circuit. A system of standing waves may be produced on such a circuit by impressing an alternating voltage of the proper frequency on the system. Under these conditions the line is in resonance with the impressed frequency. The distance along the circuit corresponding to one wave length is given by

$$D = \frac{1}{f\sqrt{LC}} \tag{5}$$

where L and C are the inductance and capacitance per unit length of the circuit and f is the frequency of the applied voltage.

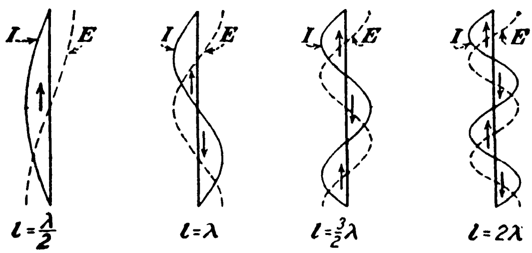


FIG. 276.—Current and voltage distributions of ungrounded antennas.

The fundamental natural frequency of oscillation for a wire in space is $c/2l$, where c is the velocity of light and l is the length of the wire. The wire will also have other natural periods corresponding to all multiples of the fundamental frequency. This is shown in Fig. 276. The current distribution will always be such as to produce a node of current at the free ends of the wire. The use of an ungrounded conductor of this type as an antenna is usually confined to the field of short waves. The conductor may be either vertical or horizontal.

The current and voltage distribution in the case of a vertical antenna when the lower end is grounded is shown in Fig. 277. By adjusting either the length of the wire or the frequency of the applied voltage to obtain resonance, a stationary wave of current is produced, as in (a), which is everywhere in phase. The current is a maximum at the base and the distribution will be sinusoidal if the inductance and capacitance per unit length are the same

throughout the structure. In this case the length of the wire required will be exactly a quarter wave length. In practice the actual length of wire required to produce resonance at a given frequency is found to be less than $\lambda/4$—usually about $\lambda/4.5$—because L and C are not uniform. This causes the current and voltage to deviate from a sinusoidal distribution. The deviation of the current from its theoretical distribution is particularly pronounced where a vertical tower is used as an antenna.[1]

If the applied frequency is doubled, or, with the same frequency, the height of the antenna is doubled, the distribution of

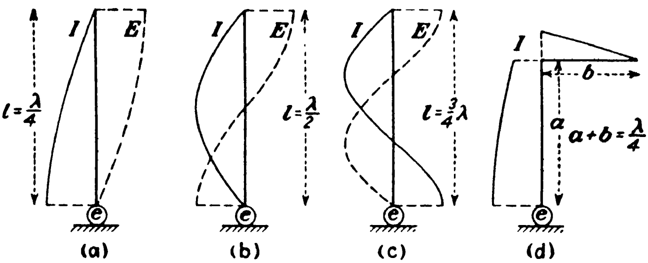


FIG. 277.—Current and voltage distributions of grounded antennas.

Fig. 277*b* is obtained. The current at the base will be zero, in the absence of any losses in the antenna, while the voltage is a maximum. In this case the structure behaves as a parallel-resonant circuit, while the previous case was one of series resonance. The load impedance offered by the antenna is consequently very high and requires an inconveniently large voltage for excitation. For this reason a grounded half-wave antenna is not commonly used for transmission purposes. A similar condition results whenever a grounded antenna is operated at any even number of quarter wave lengths.

Increasing the impressed frequency to three times the fundamental results in the distribution shown in Fig. 277*c*. Quarter-wave operation of an "inverted L" antenna is shown in (*d*).

These various modes of operation alter the vertical radiation pattern around the antenna and enable the maximum radiation to be directed toward the horizon, or at various angles of elevation, as will be explained later.

[1] H. E. GIHRING and G. H. BROWN, General Considerations of Tower Antennas for Broadcast Use, *Proc. I.R.E.*, vol. 23, p. 311, April, 1935.

It is possible to secure resonant operation of an antenna at any fraction of its fundamental wave length by connecting a condenser in series, as shown in Fig. 278*a*. The condenser annuls the inductive reactance possessed by the antenna at frequencies above its fundamental and in effect reduces its length. Most broadcasting stations operate in this manner so as to produce maximum radiation in the horizontal direction. If an inductance is placed in series with the antenna, as in Fig. 278*b*, resonant operation may be obtained at wave lengths longer than the fundamental. Adding inductance increases the effective length of the antenna.

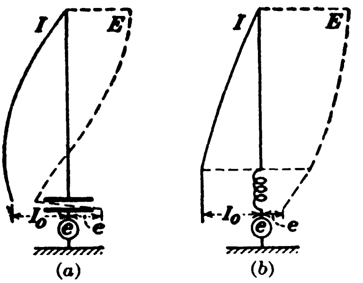


FIG. 278.—Effect of capacitance or inductance in series with an antenna.

170. Electrical Images.—If an antenna is located over a perfectly conducting earth, the electromagnetic waves radiated in the direction of the latter will be reflected from the earth's surface in the manner shown in Fig. 279. The reflected wave can be regarded as having been radiated from the electrical image of the antenna which is carrying the same current and oscillating in the same phase. This image is located in the identical position of an optical image if the earth's surface were a mirror. The effect of the earth can accordingly be determined in the case of either a grounded or ungrounded antenna by replacing the earth with an image antenna.

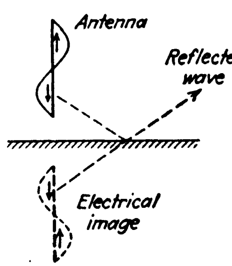


FIG. 279.—Electrical image of an antenna.

The finite conductivity of the earth may be taken into account by modifying the magnitude and phase of the current in the image in accordance with the actual values of conductivity and dielectric constant.[2] This correction is usually unnecessary at broadcast wave lengths, but at the shorter waves the error becomes appreciable for radiation at low angles. In this case the earth can be regarded as a dielectric, which can be taken into account in the image antenna of Fig. 279 by assuming the magni-

[2] T. L. ECKERSLEY, Short-wave Wireless Telegraphy, *Jour. I.E.E.* (London), vol. 65, p. 600, June, 1927.

tude of the current to be the same in both antennas but of opposite phase.

171. Electromagnetic Radiation.—The magnetic potential at a point P distant r cm. from a point pole of strength m is[3]

$$U = \frac{m}{r}$$

Differentiating U with respect to r,

$$\frac{\partial U}{\partial r} = -\frac{m}{r^2}$$

Since m/r^2 is the magnetizing force H at the point P due to the pole m,

$$H = -\frac{\partial U}{\partial r} \tag{6}$$

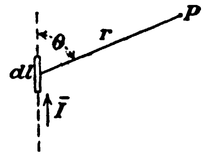


FIG. 280.—Notation for Eq. (7).

In a similar manner in Fig. 280 the magnetic potential at a point P due to a current of $\bar{I}$ abamperes flowing in the elementary length of wire dl is

$$dU = -\frac{\bar{I}dl}{r}\sin\theta \tag{7}$$

The magnetizing force at the point P, from (6), will be

$$dH = -\frac{\partial(dU)}{\partial r} \tag{8}$$

If the current in the wire at any instant is given by

$$i = \bar{I}_m \sin \omega t$$

a disturbance originating at dl at the time t will arrive at point P at a time r/c sec. later, where c is the velocity of light. The potential at P will then be given by

$$dU = -\frac{\bar{I}_m dl \sin\theta}{r}\sin\omega\left(t - \frac{r}{c}\right) \tag{9}$$

Substituting (9) in (8),

$$dH = -\bar{I}_m dl \sin\theta\left[\frac{\sin\omega\left(t - \frac{r}{c}\right)}{r^2} + \frac{\omega\cos\omega\left(t - \frac{r}{c}\right)}{cr}\right] \tag{10}$$

[3] A. S. LANGSDORF, "Principles of Direct-current Machines," 4th ed., p. 45.

It will be observed that (10) contains two terms displaced from each other in time phase by 90 degrees. The first is the "induction field" and represents energy stored in the magnetic field surrounding the conductor. This energy is alternately stored in the field and returned to the source during each half cycle. It is this field which is responsible for the phenomenon of self-inductance and is the one with which electrical engineers are chiefly concerned. The induction field diminishes as the square of the distance r and is only appreciable in the vicinity of the wire.

The second term represents the "radiation field" and is the one depicted by the vectors H in Fig. 275. Since

$$\frac{\omega}{c} = \frac{2\pi f}{c} = \frac{2\pi}{\lambda}$$

the radiation term may be written

$$dH = -\frac{2\pi I_m dl \sin\theta}{\lambda r} \cos\omega\left(t - \frac{r}{c}\right) \tag{11}$$

where the wave length λ is in centimeters.

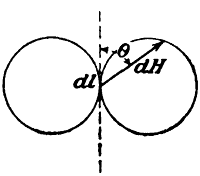


FIG. 281.—Polar diagram of relative field intensity around a radiating element.

This term represents energy propagated outward from the conductor with the velocity of light and which is not returned to the source. The radiation term diminishes as the first power of the distance r and varies inversely with the wave length.

The distribution of the radiation around the axis of the element is proportional to sin θ, resulting in a plane polar diagram resembling a figure eight, as shown in Fig. 281.

172. Field Distribution of Vertical Quarter-wave Antenna.—The field distribution around a vertical quarter-wave antenna may be obtained by determining the strength of the field at any point P due to the current in an elementary length dl and integrating over the length of the antenna. Assuming the current to be distributed sinusoidally along the conductor, the amplitude of the current at any point a distance l above the ground in Fig. 282 will be

$$I = I_0 \cos\frac{2\pi l}{\lambda} \tag{12}$$

where I_0 is the maximum amplitude of the current at the base of the antenna in amperes.

The field strength at a point P sufficiently remote from the element dl so that r_1 and r_0 may be regarded as parallel lines, from (11), is

$$dH = \frac{-0.2\pi}{\lambda r_1} I_0 \cos \frac{2\pi l}{\lambda} dl \sin \theta \cdot \cos \omega\left(t - \frac{r_1}{c}\right) \tag{13}$$

The factor 2 in (11) becomes 0.2 in (13), since I_0 is expressed in amperes instead of abamperes. If the distance to point P is

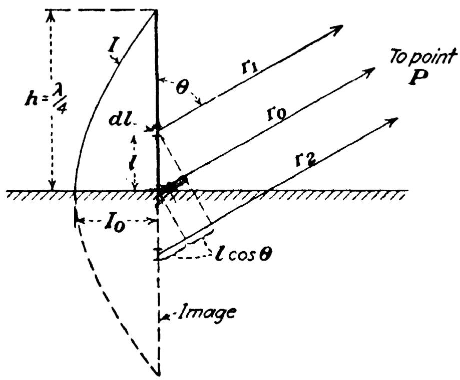


FIG. 282.—Notation for determination of field distribution of quarter-wave antenna.

large compared to l, r_1 may be replaced by r_0 in the denominator of (13), but not in the phase term $\omega\left(t - \frac{r_1}{c}\right)$, where

$$r_1 = r_0 - l \cos \theta$$

Equation (13) then becomes

$$dH = \frac{-0.2\pi}{\lambda r_0} I_0 \cos \frac{2\pi l}{\lambda} dl \sin \theta \cdot \cos \omega\left(t - \frac{r_0 - l \cos \theta}{c}\right) \tag{14}$$

Assuming a perfectly conducting earth, the reflection from the ground may be taken into account by adding on to (14) the radiation of a similar element located below the earth a distance $(-l)$. The total field strength at P will then be

$$dH = \frac{-0.2\pi I_0}{\lambda r_0} \cos \frac{2\pi l}{\lambda} \cdot \sin \theta \left[\cos \omega\left(t - \frac{r_0 + l \cos \theta}{c}\right) + \cos \omega\left(t - \frac{r_0 - l \cos \theta}{c}\right) \right] dl \quad (15)$$

From trigonometry

$$\cos (A + B) + \cos (A - B) = 2 \cos A \cos B$$

where $A = \omega\left(t - \frac{r_0}{c}\right)$ and $B = \frac{\omega l \cos \theta}{c} = \frac{2\pi l \cos \theta}{\lambda}$, so that

$$H = \frac{-0.4\pi I_0}{\lambda r_0} \sin \theta \cdot \cos \omega\left(t - \frac{r_0}{c}\right) \int_0^{\frac{\lambda}{4}} \cos \frac{2\pi l}{\lambda} \cdot \cos \frac{2\pi l \cos \theta}{\lambda} dl \quad (16)$$

The definite integral is of the form

$$\int_0^{\frac{\lambda}{4}} \cos mx \cdot \cos nx dx = \left[\frac{\sin (m - n)x}{2(m - n)} + \frac{\sin (m + n)x}{2(m + n)} \right]_0^{\frac{\lambda}{4}} \quad (17)$$

where $m = \frac{2\pi}{\lambda}$ and $n = \frac{2\pi}{\lambda} \cos \theta$. Substituting these values in (17), we get for the value of the definite integral

$$\frac{\lambda \cos \left(\frac{\pi}{2} \cos \theta\right)}{2\pi \sin^2 \theta}$$

and (16) becomes

$$H = \frac{-0.2 I_0 \cos \left(\frac{\pi}{2} \cos \theta\right)}{r_0 \sin \theta} \cos \omega\left(t - \frac{r_0}{c}\right) \quad (18)$$

which is the strength of the magnetic field (since $B = H$ in air) at a distance r_0 cm. away from a quarter-wave antenna, provided r_0 is large compared to λ. At greater distances the effect of attenuation caused by the propagation of these waves over an imperfect earth begins to be a factor and the expression becomes increasingly inaccurate.

From (3), $\mathcal{E} = 300H$, so that the field strength in volts per centimeter at any instant will be

$$\mathcal{E} = \frac{-60 I_0 \cos\left(\frac{\pi}{2}\cos\theta\right)}{r_0 \sin\theta} \cos\omega\left(t - \frac{r_0}{c}\right) \tag{19}$$

The maximum amplitude of $\mathcal{E}$ will be

$$\mathcal{E}_m = \frac{60 I_0 \cos\left(\frac{\pi}{2}\cos\theta\right)}{r_0 \sin\theta} \text{ volts per centimeter} \tag{20}$$

if I_0 is the maximum value of the current in amperes at the base of the antenna. If the effective or r.m.s. value of antenna current is used, (20) will be the effective value of the field strength.

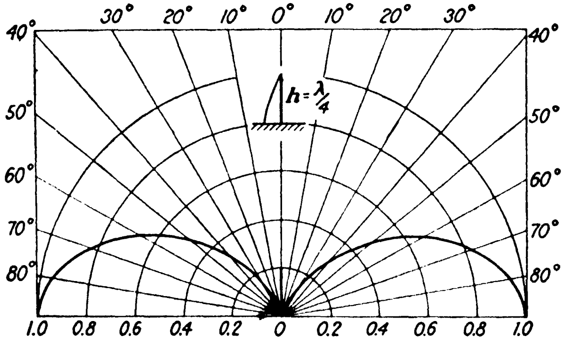


FIG. 283.—Distribution of energy in a vertical plane around a quarter-wave antenna.

The latter will have its greatest value along the horizon where $\theta = 90$ degrees.

An electromagnetic or an electrostatic field represents a definite amount of energy per unit volume of the field. In the case of air this energy is given by

$$W = \frac{\mathcal{E}^2}{8\pi} = \frac{H^2}{8\pi} \tag{21}$$

in ergs per cubic centimeter if $\mathcal{E}$ is in e.s.u. volts per centimeter. If $\mathcal{E}$ is the maximum amplitude, the average energy will be $\mathcal{E}_m^2/4\pi$.* The relative distribution of the energy around a quarter-wave antenna may then be determined by squaring the value of $\mathcal{E}_m$ in (20). This is shown in Fig. 283. The length of a radius

* This assumes $\mathcal{E}$ to be a sine function. The average value of a sine squared wave over a complete cycle is one-half the maximum value.

vector drawn from the origin to a point on the curve is proportional to the energy radiated in that direction.

The intensity of the radiation along the surface of the earth, spoken of as the "ground wave," governs the primary service area of a broadcasting station. The energy radiated at the higher angles above the horizontal is reflected from the ionosphere back to the earth again, where it interferes with the ground wave and produces fading. While it is this "sky wave" which is responsible for reception at great distances, at moderate distances where both waves are of comparable intensities the fading may be quite serious, resulting in very unsatisfactory reception. Consequently, with a given amount of power supplied to the antenna it is desirable for the radiated energy to be confined to the horizontal direction and to keep the high-angle radiation as small as possible so as to reduce the sky wave. This may be accomplished by operating at a wave length shorter than the fundamental, as will be discussed in the following section.

173. Field Distribution around Vertical Antenna.—Broadcast-transmitting antennas are at present nearly always operated at wave lengths below their fundamentals, although the established practice formerly was to operate at a wave length above the fundamental. The derivation of an expression for the field strength in the vicinity of an antenna which is not operated at an integral number of quarter wave lengths is similar to the preceding case of a quarter-wave antenna.

Assuming the current distribution to be sinusoidal and the antenna to be operated at a wave length below its natural wave length, as in Fig. 284, the amplitude of the current at a height l above the earth will be

$$i = I_m \sin \frac{2\pi(h - l)}{\lambda} = I_0 \frac{\sin \dfrac{2\pi(h - l)}{\lambda}}{\sin \dfrac{2\pi h}{\lambda}} \tag{22}$$

where I_m is the maximum value or amplitude of the current at an antinode, and I_0 is the amplitude of the current at the base of the antenna.

The radiation field, including the reflection from the ground, at a point P sufficiently remote so that r_1, r_0, and r_2 may be regarded as parallel lines, will be the same as given by (16) if we

substitute the value of i given by (22) in place of the value given by (12). Accordingly,

$$H = \frac{-0.4\pi I_m}{\lambda r_0} \sin\theta \cdot \cos\omega\left(t - \frac{r_0}{c}\right) \int_0^h \sin\frac{2\pi(h-l)}{\lambda} \cdot \cos\frac{2\pi l\cos\theta}{\lambda} dl \quad (23)$$

The value of the definite integral is

$$\int_0^h \sin\frac{2\pi(h-l)}{\lambda} \cos\frac{2\pi l\cos\theta}{\lambda} dl = \frac{\cos\frac{2\pi h\cos\theta}{\lambda} - \cos\frac{2\pi h}{\lambda}}{\frac{2\pi}{\lambda}\sin^2\theta} \quad (24)$$

Substituting (24) in (23),

$$H = \frac{0.2I_m}{r_0\sin\theta}\left(\cos\frac{2\pi h}{\lambda} - \cos\frac{2\pi h\cos\theta}{\lambda}\right)\cos\omega\left(t - \frac{r_0}{c}\right) \quad (25)$$

which reduces to the value given by (18) when $h = \lambda/4$.

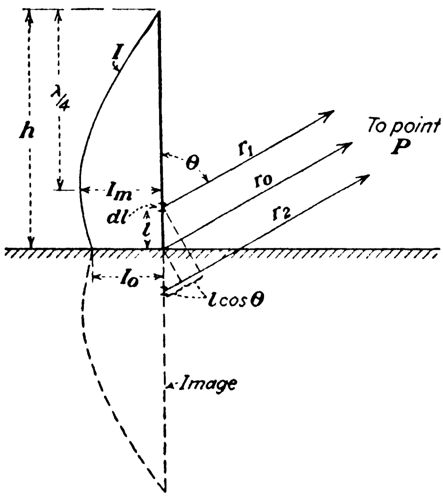


FIG. 284.—Notation for determination of field strength around grounded vertical antenna operated below natural wave length.

From (3), the strength of the electrostatic field in volts per centimeter will be

$$\mathcal{E} = \frac{60I_m}{r_0\sin\theta}\left(\cos\frac{2\pi h}{\lambda} - \cos\frac{2\pi h\cos\theta}{\lambda}\right)\cos\omega\left(t - \frac{r_0}{c}\right) \quad (26)$$

where r_0 is the distance from the antenna in centimeters and c is the velocity of light in centimeters per second.

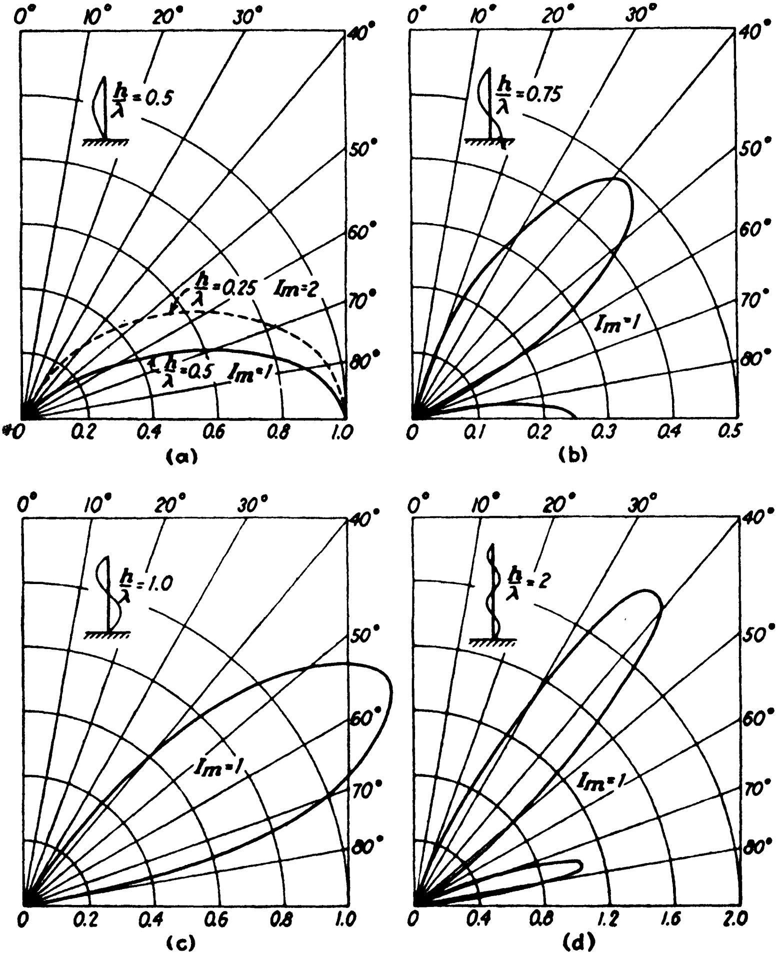


FIG. 285.—Distribution of energy in vertical plane for grounded vertical antennas.

The maximum amplitude of $\mathcal{E}$ will be

$$\mathcal{E}_m = \frac{60 I_m}{r_0 \sin \theta}\left(\cos \frac{2\pi h}{\lambda} - \cos \frac{2\pi h \cos \theta}{\lambda}\right) \text{ volts per centimeter} \qquad (27)$$

where I_m is the amplitude of the antenna current at an antinode and is related to the current at the base of the antenna by (22).

The distribution of the radiated energy in the vertical plane will be proportional to the square of (27), and is shown in Fig. 285 for various ratios of antenna height to operating wave length. As the height of the antenna is increased, or the operating wave length is reduced by the insertion of a series condenser at the base to maintain resonance, the radiation diminishes at angles above the horizontal, as shown in (*a*). The dotted curve for the quarter-wave antenna of Fig. 283 is shown for comparison. The curve becomes progressively flatter as the ratio of h/λ is increased. The area under the curve is proportional to the power being radiated and it is evident that the half-wave antenna will produce the same field intensity along the horizon with less power supplied to the antenna than with quarter-wave operation. Stated another way, with a constant amount of power supplied to the antenna the strength of the ground wave increases as the ratio of h/λ is made larger, the optimum ratio being $h/\lambda = 0.625$.

When h/λ becomes greater than 0.5 the radiated energy breaks up into two parts and a secondary lobe appears at high angles. The two lobes are separated by a conical surface along which the radiation is zero, in addition to the dead spot directly above the antenna. The size of the secondary lobe increases, while that of the primary lobe diminishes, the two having the areas shown in Fig. 285*b* when $h/\lambda = 0.75$. When $h/\lambda = 1$, the primary lobe disappears entirely and the maximum energy is now directed along a line about 30 degrees above the horizontal. Under this condition the ground wave would be zero, theoretically, and all of the radiated energy would be concentrated in the sky wave. As h/λ is made greater than unity, two lobes again appear, resembling the curves of Fig. 285*b* when $h/\lambda = 1.5$. The polar distribution of the energy for $h/\lambda = 2$ is shown in Fig. 285*d*.

The distribution about an inverted L antenna operated at its natural period has been shown by G. W. Pierce[4] to be

$$\mathcal{E}_m = \frac{60 I_m}{r_0 \sin \theta}[\cos B \cos (A \cos \theta) - \sin B \cos \theta \sin (A \cos \theta) - \cos (A + B)] \quad (28)$$

where $A = 2\pi a/\lambda$
$B = 2\pi b/\lambda$
a = vertical height.
b = horizontal length.

[4] Electric Oscillations and Electric Waves, Chap. IX, p. 433.

The expression for the vertical distribution in the case of an ungrounded vertical antenna located a distance d above the earth and operating at an integral number of half wave lengths, may be obtained in a manner similar to the derivation for the grounded antenna. The field intensity in volts per centimeter at an angle θ from the axis of the antenna when operating at an odd number of half wave lengths is given by[5]

$$\mathcal{E}_m = \frac{60I_m}{r_0 \sin \theta} \cos \left[\frac{\pi}{2}n\left(1 + \frac{2d}{l}\right) \cos \theta\right] \cos \left(\frac{\pi}{2}n \cos \theta\right) \quad (29)$$

where I_m = current at an antinode, amperes.
r_0 = distance from antenna, centimeters.
$n = 2l/\lambda$, the number of half wave lengths (odd).
l = length of antenna.
d = distance of lower end of antenna above the earth.

When n is an even number of half wave lengths the term $\cos\left(\frac{\pi}{2}n \cos \theta\right)$ in (29) becomes $\sin\left(\frac{\pi}{2}n \cos \theta\right)$. The polar distribution curves for two typical cases are shown in Fig. 286.

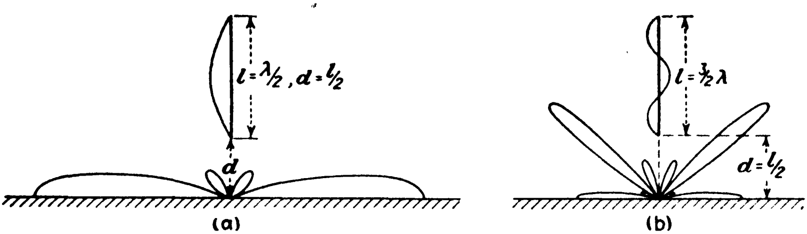


FIG. 286.—Distribution of energy in vertical plane for ungrounded vertical antennas.

Energy for the excitation of these antennas is supplied by a suitable nonradiating transmission line which is coupled to the antenna.

If the antenna is sufficiently remote so that the reflection from the ground can be neglected, (29) becomes[6]

[5] S. A. LEVIN and C. J. YOUNG, Field Distribution and Radiation Resistance of a Straight Vertical Unloaded Antenna Radiating at One of Its Harmonics, *Proc. I.R.E.*, vol. 14, p. 675, October, 1926.

[6] P. S. CARTER, C. W. HANSELL, and N. E. LINDENBLAD, Development of Directive Transmitting Antennas by RCA Communications, Inc., *Proc. I.R.E.*, vol. 19, p. 1773, October, 1931.

$$\mathcal{E}_m = \frac{60I_m}{r_0 \sin\theta} \cos\left(\frac{\pi}{2} n \cos\theta\right) \tag{30}$$

when the length of the antenna is an odd number, n, of half waves long, and

$$\mathcal{E}_m = \frac{60I_m}{r_0 \sin\theta} \sin\left(\frac{\pi}{2} n \cos\theta\right) \tag{31}$$

when n is an even number of half wave lengths. Both expressions are in volts per centimeter.

Typical distribution curves are shown in Fig. 287, which are cross sections of a figure of revolution in three-dimensional space. As the length of the antenna measured in half waves increases, the number of lobes is increased, the number being equal to the length in half waves. As will be observed, the radiation pattern consists of two principal lobes and a number of minor ones when the wire is several waves in length. By using a group of wires properly spaced and fed in the proper phase an antenna is obtained which is highly directional in both the vertical and horizontal planes. Such an arrangement is called an *array*. Directional characteristics obtained in this manner are extensively used in short waves where the physical lengths and spacings required result in structures of reasonable size.

The directional characteristics of all the foregoing types of antennas apply equally to reception as well as transmission. The characteristics of an antenna when used to abstract energy from a passing electromagnetic wave are similar in practically all respects to the corresponding characteristics of the same structure when acting as a radiator of energy.[7]

174. Radiation Resistance.—The radiation of energy by an antenna may be accounted for by assuming it to be equivalent to a fictitious I^2R loss. This loss is in addition to the dielectric losses associated with the antenna and the ohmic losses in the antenna conductors. The radiation resistance is defined as

$$R_r = \frac{P}{I^2} \tag{32}$$

[7] STUART BALLANTINE, Reciprocity in Electromagnetic, Mechanical, Acoustical, and Interconnected Systems, p. 929; and J. R. CARSON, Reciprocal Theorems in Radio Communication, p. 952, *Proc. I.R.E.*, vol. 17, June, 1929.

where P is the power actually radiated and I is the r.m.s. value of the antenna current. Since the distribution of current varies continuously along the length of the antenna, the value of I used

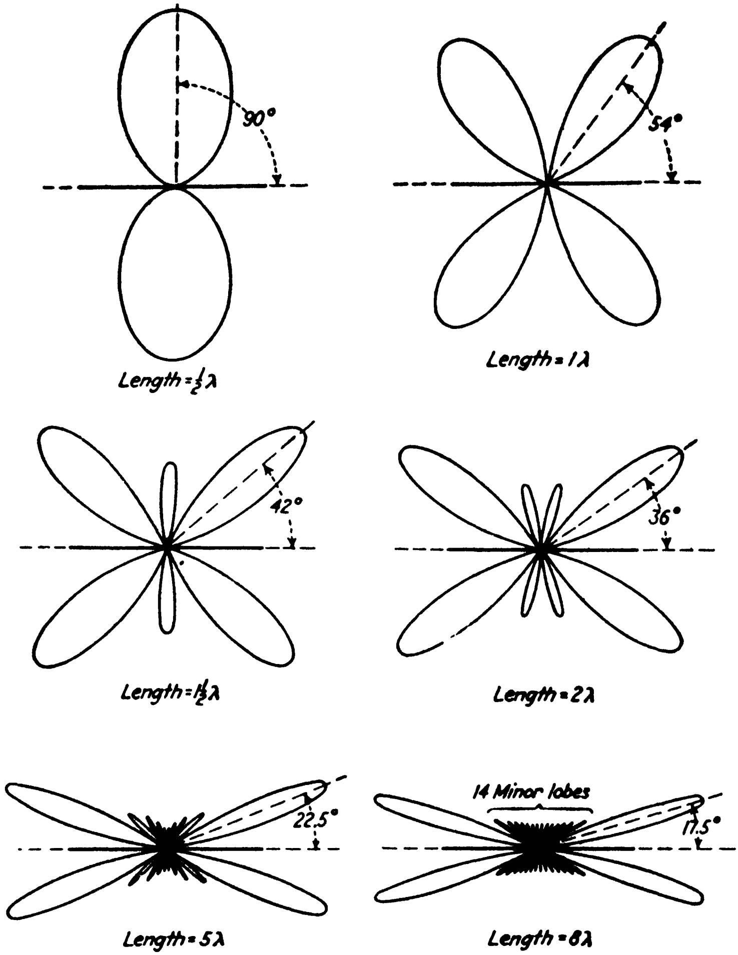


FIG. 287.—Polar distribution of energy around an antenna remote from the earth.

in (32) is usually that flowing at the point where the antenna is fed. This may be either at the base or at a current antinode.

The radiation resistance of a quarter-wave antenna may be determined by summing up the density of energy due to the electromagnetic (or electrostatic) field in a hemispherical shell

above the earth's surface. The area of an elementary zone on the surface of a hemisphere of radius r_0 will be

$$da = 2\pi r_0^2 \sin\theta d\theta$$

In a shell 1 cm. thick the elementary volume will be

$$dv = 2\pi r_0^2 \sin\theta d\theta$$

The total energy per unit volume is $H^2/4\pi$, so that the energy in the hemispherical shell will be

$$W = \int_0^{\frac{\pi}{2}} \frac{H^2}{4\pi} dv = \frac{1}{2}\int_0^{\frac{\pi}{2}} H^2 r_0^2 \sin\theta d\theta \tag{33}$$

Substituting the value of H from (18) in (33), we get

$$W = 2\bar{I}_0^2 \cos^2 \omega\left(t - \frac{r_0}{c}\right) \int_0^{\frac{\pi}{2}} \frac{\cos^2\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} d\theta \tag{34}$$

where $\bar{I}_0$ is in abamperes. The value of the definite integral in (34) is

$$\int_0^{\frac{\pi}{2}} \frac{\cos^2\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} d\theta = 0.6095$$

and (34) becomes

$$W = 1.219\bar{I}_0^2 \cos^2 \omega\left(t - \frac{r_0}{c}\right) \tag{35}$$

The average value of $\cos^2 \omega\left(t - \frac{r_0}{c}\right)$ over a complete cycle is one-half, so that

$$W_{av} = 0.6095\bar{I}_0^2 \tag{36}$$

This energy is flowing outward through the shell with the velocity of light c and the power radiated will be

$$P = cW_{av} = 1.8285 \times 10^{10}\bar{I}_0^2 \tag{37}$$

From (32),

$$R_r = \frac{P}{\bar{I}^2} = \frac{2P}{\bar{I}_0^2} = 3.657 \times 10^{10} \text{ abohms} \tag{38}$$

since $\bar{I}_0$ is the amplitude or maximum value of the current at the base of the antenna in abamperes. The radiation resistance of a quarter-wave antenna in practical units is

$$R_r = 36.57 \text{ ohms} \tag{39}$$

This value is independent of the frequency and a vertical antenna 2.5 meters high operating at a wave length of 10 meters (30,000 kc) will have the same radiation resistance as one 75 meters high operating at 300 meters (1000 kc).

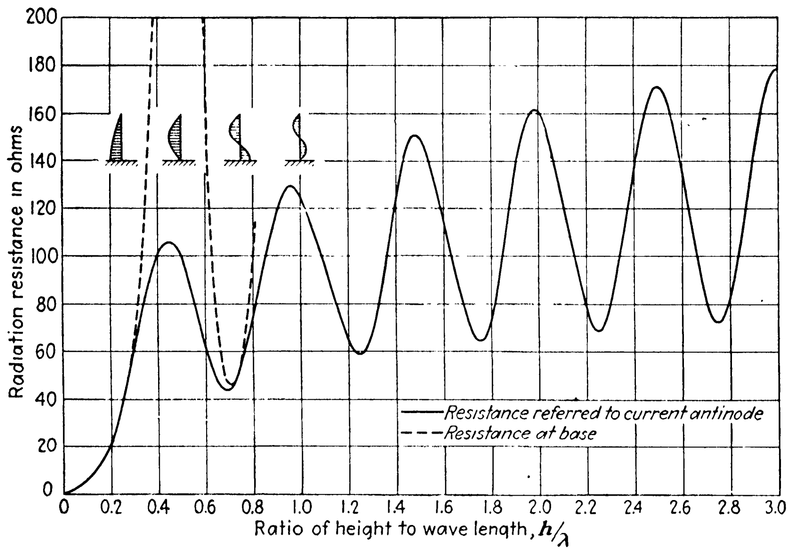


FIG. 288.—Radiation resistance of vertical grounded antenna.

The radiation resistance for various other values of h/λ may be obtained in a similar manner by substituting the appropriate expression for H in (33) and integrating over the surface of the hemisphere. The evaluation of the definite integral becomes rather complicated when the antenna is operated below the fundamental wave length. These values of radiation resistance have been computed by Ballantine[8] and are plotted in Fig. 288 as a function of h/λ, assuming a perfectly conducting earth. This resistance may be referred either to the current I_0 at the

[8] On the Radiation Resistance of a Simple Vertical Antenna at Wave Lengths below the Fundamental, *Proc. I.R.E.*, vol. 12, p. 823, December, 1924. See also discussion, *Proc. I.R.E.*, vol. 15, p. 245, March, 1927.

base of the antenna, as shown by the dotted curve, or to the current I_m at an antinode. These two currents are the same when the antenna is operated at an odd number of quarter wave lengths. When operated at an even number of quarter wave lengths, the current at the base is theoretically zero and the radiation resistance referred to I_0 is infinite.

The actual measured resistance of an antenna will include, in addition to the radiation resistance, the ohmic resistance of the wires and ground connection, and also the resistance due to an imperfect dielectric surrounding the antenna caused by buildings, trees, etc. At long wave lengths where the antenna is apt to be operated considerably below its fundamental, because of the physical limitations as to its size, the radiation resistance may be only a small portion of the total. The radiation efficiency in such cases will be comparatively low. At short waves the radiation resistance is a large percentage of the total, so that with a given amount of power supplied to an antenna the major portion will be radiated.

When the antenna is located over ground that is an extremely poor conductor, the ground connection is dispensed with entirely and the connection is made to an insulated counterpoise instead. This consists of a network of wires suspended above the earth and extending outward from the base of the antenna. The area of the counterpoise should be at least equal to, and preferably greater in extent than, the flat-topped portion in the case of an inverted L or T antenna. In the case of a vertical antenna the wires of the counterpoise should extend radially outward to a distance at least equal to the antenna height. With soil of good conductivity a similar structure of wires is usually buried under the surface of the earth to insure a ground connection of low resistance.

175. Wave Length for Optimum Ground Wave.—Maximum field strength in the horizontal direction for a given amount of radiated power is obtained by operating the antenna below its fundamental wave length.[9] Where the ground wave is the important factor, as in broadcasting, it is desirable to keep the energy radiated at high angles at a minimum.

[9] STUART BALLANTINE, On the Optimum Transmitting Wave Length for a Vertical Antenna over Perfect Earth, *Proc. I.R.E.*, vol. 12, p. 833, December, 1924.

The amplitude of the field intensity in the horizontal direction where $\theta = 90$ degrees, from (27), is

$$\mathcal{E} = \frac{60 I_m}{r_0}\left(\cos\frac{2\pi h}{\lambda} - 1\right) \tag{40}$$

The radiated power is

$$P = I_m^2 R_r \tag{41}$$

where R_r is referred to the current at an antinode. From (41)

$$I_m = \sqrt{\frac{P}{R_r}}$$

and substituting this value for I_m in (40), we get

$$\mathcal{E} = \frac{60}{r_0}\sqrt{\frac{P}{R_r}}\left(\cos\frac{2\pi h}{\lambda} - 1\right)$$

or

$$\frac{\mathcal{E}}{\sqrt{P}} = \frac{60}{r_0}\,\frac{\cos\frac{2\pi h}{\lambda} - 1}{\sqrt{R_r}} \tag{42}$$

For a given amount of radiated power, we are interested in making $\mathcal{E}/\sqrt{P}$ a maximum. The variation of R_r as a function of h/λ is given in Fig. 288, and from these data the variation of $\mathcal{E}/\sqrt{P}$ may be computed. The results are given in Fig. 289. It will be observed that the optimum horizontal field strength is obtained for a value of $h/\lambda = 0.625 = 5/8$. The distribution of energy in the vertical plane will be similar in appearance to Fig. 286*a*, the maximum of the secondary lobe being directed at an angle of about 55 degrees above the horizontal. This lobe of high angle radiation will be reflected from the ionosphere back to earth where it interferes with the ground wave and causes fading at distances where the intensities of the two waves are comparable in magnitude. In order to increase the incipient fading distance and thereby increase the area of satisfactory reception, broadcasting stations employing vertical radiators frequently use a value of h/λ somewhat less than optimum which reduces the size of the secondary lobe.[10] When the current in the antenna

[10] STUART BALLANTINE, High Quality Radio Broadcast Transmission and Reception. *Proc. I.R.E.*, vol. 22, p. 564, May, 1934.

deviates appreciably from a sinusoidal distribution, this secondary lobe may be entirely absent. Such deviations may be caused by variations in the values of L and C per unit length, as when a tapered vertical tower, which is insulated from the ground at the base, is used as a radiator. The finite conductivity of the earth is also a factor, causing the minimum between the two lobes to be less sharply defined.

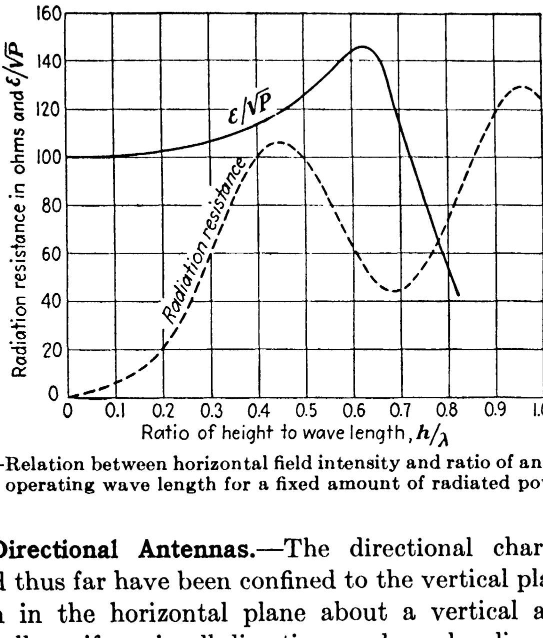


FIG. 289.—Relation between horizontal field intensity and ratio of antenna height to operating wave length for a fixed amount of radiated power.

176. Directional Antennas.—The directional characteristics discussed thus far have been confined to the vertical plane. The radiation in the horizontal plane about a vertical antenna is theoretically uniform in all directions and a polar diagram of the energy distribution would be a circle with the antenna at the center. Practically, the character of the terrain, the presence of buildings, trees, bodies of water, etc., will modify this distribution; particularly as the frequency is increased and the size of the obstacles becomes comparable to the wave length used. A uniform distribution of energy around the antenna is entirely satisfactory for broadcasting, but for point-to-point communication the energy radiated in directions other than the desired one is wasted and causes interference with other stations. A directional receiving antenna is highly desirable in point-to-point communication, in that interference and noise originating in

directions other than that giving maximum response will be greatly reduced.

One method of securing directional characteristics in the horizontal plane is to mount a vertical antenna along the focal line of a parabolic cylinder of sheet metal, as illustrated in Fig. 290. Increasing the size of the reflector increases the concentration of the beam as shown. The sheet-metal surface may be replaced by a series of spaced vertical wires one-half wave length long and parallel to each other, arranged in the form of a parabolic cylinder. This structure is just as satisfactory as a continuous metal sheet, if the individual spacing is not greater than a quarter

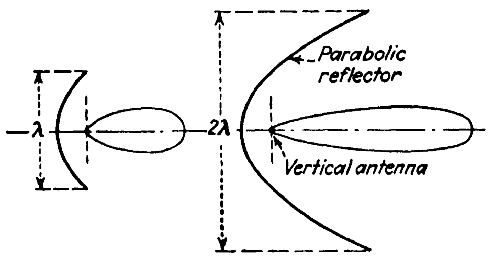


FIG. 290.—Concentration of radiation in horizontal plane by means of parabolic reflectors.

wave length. Reflectors of this type were used in the original experiments of Hertz and the advantages to be gained by concentrating the radiated energy into a beam were appreciated by a number of early investiga . However, the early trend of development was toward the use of the longer wave lengths which rendered directional antennas impractical because of their prohibitive size. The scheme of Fig. 290 has been used in some of the earlier short-wave transmitters when this mode of transmission began to replace long-wave systems. Newer directive systems have been developed of simpler mechanical construction and greater gain in field strength, so that parabolic reflectors are used only for ultra-short waves.

177. Loop Antennas.—The loop or coil antenna is another directional type which has been widely used with portable receiving sets and in other special applications, such as field-strength measurements and direction finding in radio-compass stations. Closed loops are relatively inefficient radiators unless their dimensions are made comparable to the wave lengths employed, which

restricts their use for transmission. The principal application for transmission purposes is in connection with radio beacons for aviation.

If a rectangular loop of wire of height h and width w, as shown in Fig. 291, is acted upon by an electromagnetic wave of $\mathcal{E}$ volts per meter, the e.m.f. induced will be the line integral of $\mathcal{E}$ around the loop. If $\mathcal{E}$ at any instant is given by

$$\mathcal{E} = \mathcal{E}_m \sin \omega\left(t - \frac{r}{c}\right)$$

the voltages induced in the two sides of the loop will be in the same direction, but will be out of phase with each other by an

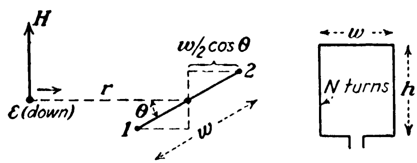


FIG. 291.—Notation for determination of voltage induced in loop antenna.

amount depending on the width w. The field strength at 1 will be

$$\mathcal{E}_1 = \mathcal{E}_m \sin \omega\left(t - \frac{r - \frac{w}{2}\cos\theta}{c}\right) \tag{43}$$

while that at 2 is

$$\mathcal{E}_2 = \mathcal{E}_m \sin \omega\left(t - \frac{r + \frac{w}{2}\cos\theta}{c}\right) \tag{44}$$

The resultant e.m.f. e induced in the loop will be

$$e = \mathcal{E}_1 h - \mathcal{E}_2 h \tag{45}$$

and if there are N turns on the coil, the instantaneous e.m.f. will be

$$e = Nh\mathcal{E}_m\left[\sin \omega\left(t - \frac{r - \frac{w}{2}\cos\theta}{c}\right) - \sin \omega\left(t - \frac{r + \frac{w}{2}\cos\theta}{c}\right)\right] \tag{46}$$

From trigonometry,

$$\sin A - \sin B = 2 \cos \tfrac{1}{2}(A + B) \sin \tfrac{1}{2}(A - B)$$

and (46) becomes

$$e = 2Nh\mathcal{E}_m \cos \frac{1}{2}\left(2\omega t - \frac{2\omega r}{c}\right) \sin \frac{1}{2}\left(\frac{\omega w \cos \theta}{c}\right) \tag{47}$$

For small values of $\frac{1}{2}\left(\frac{\omega w \cos \theta}{c}\right)$, the sine may be taken equal to the angle, and (47) becomes

$$\begin{aligned} e &= \frac{Nhw\omega \cos \theta}{c}\mathcal{E}_m \cos \omega\left(t - \frac{r}{c}\right) \\ &= \frac{NA\omega \cos \theta}{c}\mathcal{E}_m \cos \omega\left(t - \frac{r}{c}\right) \end{aligned} \tag{48}$$

where A is the area of the loop. Substituting the wave length λ for the radian frequency ω, (48) becomes

$$e = \frac{2\pi NA \cos \theta}{\lambda}\mathcal{E}_m \cos \omega\left(t - \frac{r}{c}\right) \tag{49}$$

The induced e.m.f. will be a maximum when $\cos \omega\left(t - \frac{r}{c}\right) = 1$, and is

$$e_m = \frac{2\pi NA \cos \theta}{\lambda}\mathcal{E}_m \tag{50}$$

where e_m = maximum value of induced e.m.f., volts.
N = number of turns on loop.
A = area of loop, square meters.
$\mathcal{E}_m$ = maximum field strength, volts per meter.
λ = wave length, meters.
θ = angle between plane of loop and direction of transmitting station.

The fraction multiplying $\mathcal{E}_m$ in (50) is termed the "effective height" of the loop since this factor when multiplied by the field strength in volts per meter gives the net voltage induced. The greatest value of induced voltage will occur when $\theta = 0$, or when the plane of the coil lies along the direction of the transmitted

signal. The directional characteristics are shown in Fig. 292, and apply to loops of all shapes, since the voltage induced is proportional to the area of the loop. When used as a direction finder the loop is rotated until the signal is a minimum, which occurs when $\theta = 90$ degrees. The position of the minimum signal is much more sharply defined than the maximum.

A variable condenser shunted across the loop serves as a tuning means. The voltage across the condenser at resonance is given by

$$E = \frac{i}{\omega C} = i\omega L = \frac{e}{R}\omega L = eQ \quad (51)$$

where e is the induced e.m.f. and Q is the ratio of $\omega L/R$ for the loop.

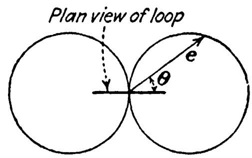


FIG. 292.—Directional characteristics of loop antenna.

178. Radio Direction Finders.[11]—If two observations are made at two receiving stations equipped with direction finders and separated from each other by a known distance, the location of a transmitting station can be readily determined by triangulation. This method is used by the government-operated radio-compass stations at the port of New York and elsewhere as an aid to navigation. Likewise, ships equipped with loop antennas can locate their positions by determining the bearings of two or more shore stations. Direction finders on board ship are also of great value in locating the position of a vessel in distress, particularly when the latter is uncertain of its exact position.

The 180-degree uncertainty in loop bearings can be eliminated by using a vertical antenna in conjunction with the loop. Referring to Fig. 293, an oncoming wave in the direction a will induce a voltage E_1 in the vertical side 1 of the loop which will be ahead in phase by an angle ϕ of E_3 induced in the opposite side of the loop. The resultant e.m.f. acting around the loop will be $e_a = E_1 - E_3$, which will be 90 degrees out of phase with E_2, the voltage induced in the vertical antenna 2 located along the vertical axis of the loop. This is shown by the vector diagram at (*a*). Had the wave been traveling in the direction b, the vector diagram would be as shown in (*b*), and the resultant volt-

[11] For additional information including an extensive bibliography, see R. L. Smith-Rose, Radio Direction-finding by Transmission and Reception, *Proc. I.R.E.*, vol. 17, p 425, March, 1929.

age e_b acting around the loop would have been just opposite in direction to e_a. The voltage induced in the vertical antenna would remain unchanged.

If the loop and the vertical antenna are coupled by means of the coils L_1 and L_2, a current in L_2 will induce a voltage of $\pm j\omega M i_2$

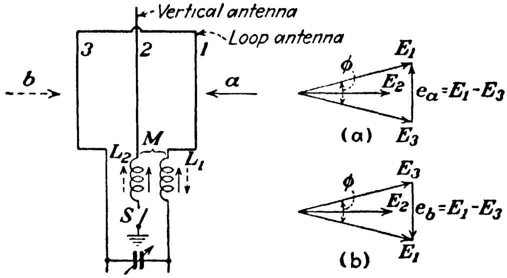


FIG. 293.—Vector diagrams of voltages induced in loop and vertical antennas, enabling direction of wave-travel to be determined.

in L_1, depending upon how the two coils are wound with respect to each other. This induced voltage, which will be 90 degrees out of phase with the current i_2 in the vertical antenna, will therefore aid the e.m.f. acting in the loop for a wave traveling in direction a and oppose the e.m.f. produced by a wave traveling in direction b. The direction of the arriving waves can then be

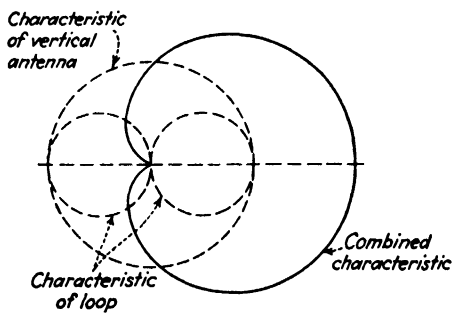


FIG. 294.—Directional characteristics of a loop in conjunction with a vertical antenna.

determined by rotating the loop for maximum signal strength ($\theta = 0$) with switch s open. With the vertical antenna tuned to the same signal, closing s will either increase or decrease the strength of the signals, depending upon the direction of propagation. The directive characteristics are shown in Fig. 294,

assuming the effective heights of the loop and vertical antenna to be the same.

In order to avoid the necessity of rotating the loop antenna, which is inconvenient when the dimensions of the latter are large, a method developed by Bellini and Tosi in 1907 consisting of two fixed loops at right angles to each other may be used. These usually consist of a single turn, triangular in shape, supported by a central mast. The outputs of the two antennas are impressed upon a goniometer consisting of two pairs of primary coils at right angles to each other, one pair for each loop, within which and free to be rotated is mounted a secondary coil, as shown in Fig. 295. For a wave traveling in a direction lying

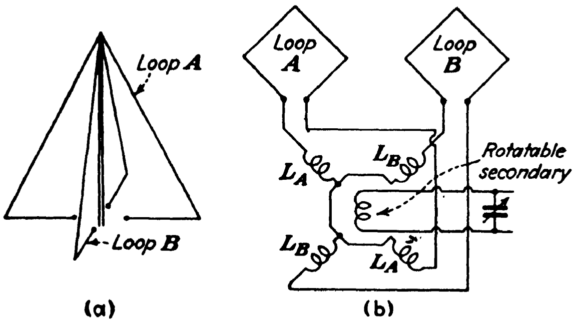


FIG. 295.—Direction finder consisting of two fixed loops at right angles to each other connected to a goniometer.

in the plane of loop A, maximum secondary voltage will be induced when the axis of the secondary coil coincides with that of L_A, while a wave traveling in the plane of loop B will produce maximum secondary voltage when the axis of the secondary coil coincides with that of L_B. A wave arriving in a direction somewhere between the planes of the two antennas will produce currents in both L_A and L_B, and the axis of the resultant magnetic field produced will be perpendicular to the direction of propagation of the arriving wave. In other words, if the planes of the coils L_A and L_B are parallel to their respective antennas, the position of the secondary coil for maximum signal strength will be exactly the same as though it were a loop antenna being acted upon directly by the incident wave. The directional characteristics of this scheme as the secondary coil is rotated will evidently be the same as Fig. 292 for the ordinary loop antenna.

In order to determine the actual direction of origin of the wave instead of merely its bearing, the characteristics of the fixed loops can be combined with that of a vertical antenna, so as to obtain the cardioid diagram of Fig. 294. In practice, the two loops themselves can be made to function as a vertical antenna as well as individual loops.[12]

This type of direction finder as well as the rotating loop should be calibrated after installation, particularly on shipboard, in order that the distorting effect of surrounding metal objects, rigging, etc., on the observed bearings may be corrected for.[13] The loop should be carefully balanced with respect to stray capacitances to ground as the "vertical-antenna effect" caused by such unbalances obscures the point of minimum response.

Another source of error in direction-finding systems is caused by down-coming sky waves which are more or less horizontally polarized. These waves induce voltages in the top and bottom portions of the loop, giving rise to a resultant e.m.f. acting around the circuit which cannot be reduced to zero by rotating the loop about a vertical axis. As a result there will either be no loop position giving zero signal, or else the position of zero signal will not be the true bearing of the transmitting station. Direction finding by either of the above systems can be relied upon only when the vertically polarized ground wave is strongly predominant. The sky wave is always very much stronger at night and bearings taken then are very much more erratic than in the daytime. The plane of polarization of the sky wave as well as the amount of attenuation it experiences may change rather rapidly, causing similar variations to occur in the radio-compass bearings. These changes in the plane of polarization are due to variations in the height and density of the ionized layer which is responsible for the reflection of the sky wave.

The deviation from the true bearing is seldom more than 2 degrees at distances of 100 miles over water in the daytime for frequencies in the neighborhood of 500 kc. For the same amount of error at night the maximum distance would be about 30 miles.

[12] H. DONISTHORPE, The Marconi Marine Radio Direction Finder, *Proc. I.R.E.*, vol. 13, p. 29, February, 1925.

[13] C. T. SOLT, The Development and Application of Marine Radio Direction Finding Equipment by the U. S. Coast Guard, *Proc. I.R.E.*, vol. 20, p. 228, February, 1932.

Observations made on the very long waves (30 kc and lower) show comparatively small deviations from the true bearings at distances of several thousand miles in the daytime. Shortly before sunset irregularities begin and continue until after sunrise. These deviations continually vary in magnitude, sometimes reaching values of 20 degrees or more. Short waves are unsuited for direction finding owing to the rapid attenuation of the ground wave.

The difficulties caused by the sky wave can be eliminated by a direction-finding system due to Adcock.[14] In its simplest form it consists of two spaced vertical antennas connected as shown in Fig. 296 and capable of being rotated about a vertical axis through the center. The vertically polarized ground wave induces voltages in the vertical members in exactly the same manner as an ordinary loop, but the connections are such as to cause the voltages induced in the horizontal members to annul each other. Since the antenna is essentially a loop of one turn, the dimensions required for reasonable sensitivity are such as to make rotation rather inconvenient. In such cases two pairs of these antennas fixed in position and at right angles to each other may be used in conjunction with a goniometer, as in the Bellini-Tosi system.

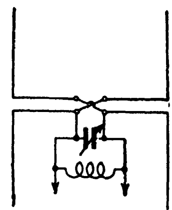


FIG. 296.—Adcock direction finder which is not affected by sky wave.

179. Radio-range Beacons.—This name is given to the directional transmitters used along established airways in the United States to enable aircraft to reach their destinations under any conditions of visibility. They are located approximately 200 miles apart, and employ either crossed loops or a modified form of the Adcock antenna. Two methods of signaling are used, the aural type and the visual type.[15]

The aural-type beacon, which was developed first, employs two crossed triangular loops of a single turn, similar to Fig. 295*a*. Each loop is individually excited from a separate power amplifier using a 500-cycle plate-voltage supply so as to produce an audible

[14] British patent No. 130,490.

[15] J. H. DELLINGER and H. PRATT, Development of Radio Aids to Air Navigation, *Proc. I.R.E.*, vol. 16, p. 890, July, 1928; also J. H. DELLINGER, H. DIAMOND, and F. W. DUNMORE, Development of the Visual-type Airway Radiobeacon System, *Proc. I.R.E.*, vol. 18, p. 796, May, 1930.

signal without the need of a heterodyne detector in the receiving set. Grid excitation for both power amplifiers is obtained from a common master oscillator. The grid circuits of the power amplifiers are alternately keyed by motor-driven cams, causing one loop to transmit the letter *A* (· —) and the other *N* (— ·); interlocked so that the dot of the *A* completely fills the space between the dash and dot of the *N*. The field strength around the two loops is shown in Fig. 297, and in the planes bisecting the angles

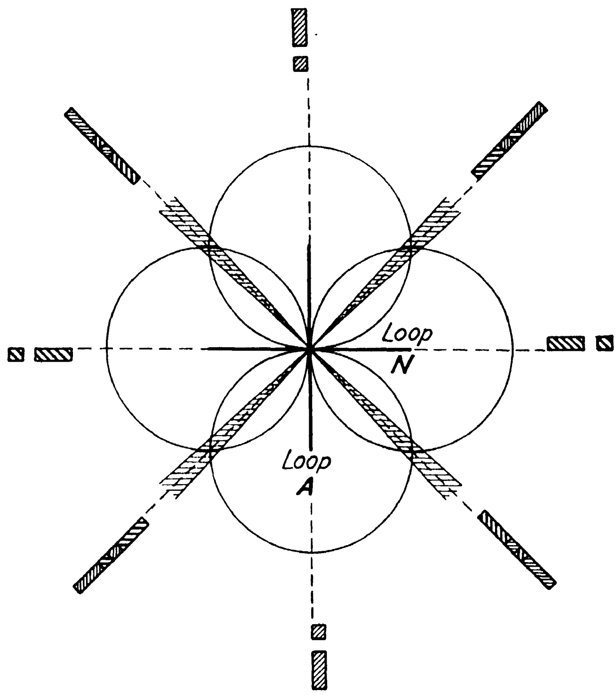


FIG. 297.—Distribution of radiation around crossed loops of aural-type radio beacon, showing equisignal zones.

between them the field strength of both antennas is the same, causing the two letters to merge together and be heard as a succession of long dashes. A deviation to the right or the left of the course will cause either the letter *A* or *N* to predominate. The width of the equisignal zone is only a few degrees.

In order to avoid rotating the loops, which are usually several hundred feet long at the base and 50 or more feet high, a suitable goniometer is used which enables the two loops to be rotated in

effect in the same manner as in the Bellini-Tosi system of directional reception. This permits the four equisignal zones to be rotated at will so as to make them coincide with the airways converging on the airport without disturbing the position of the loop antennas. The angle between the two intersecting courses can be shifted from the 90-degree relation by increasing or decreasing the current in one of the loops.

The simplest form of visual system is similar to the aural system just described. Instead of alternately exciting the separate antennas, they now radiate continuously, but are modulated at

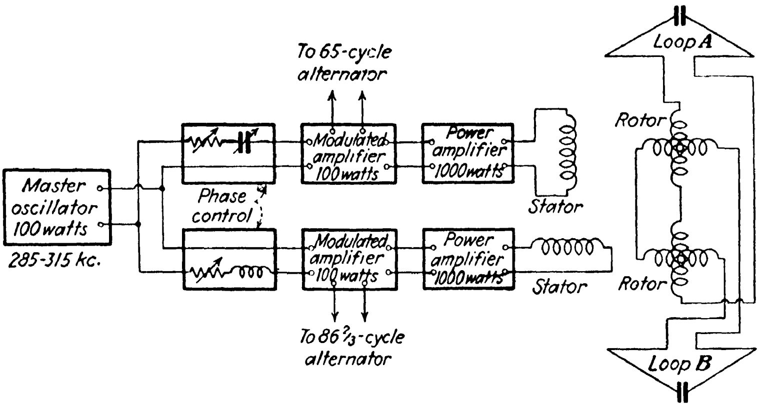


FIG. 298.—Schematic diagram of visual type of double-modulation radio beacon.

different values of audio frequency. The schematic diagram is shown in Fig. 298. If the two antennas are excited in phase—which can be accomplished by proper adjustment of the phase-control circuits ahead of the two modulated amplifiers—the radiation characteristic around each loop will consist of two figure-eight patterns at right angles to each other. The two carriers being of the same frequency and in time phase, combine to produce a resultant carrier having its maximum intensity along a plane bisecting the angle between the two loops. The side bands of the two loops do not combine, since they are of different frequencies, and therefore the maximum intensity of each side band will be in the plane of the loop from which it was radiated.[16] The resultant pattern in space is shown in Fig. 299.

[16] H. PRATT, Field Intensity Characteristics of Double Modulation Type of Directive Radio Beacon, *Proc. I.R.E.*, vol. 17, p. 873, May, 1929.

The audio-frequency output of the receiving set on the aircraft actuates a pair of tuned reeds which are resonant to the modulating frequencies used, the latter being 65 and 86⅔ cycles. These reeds, which are similar to those of a Frahm frequency meter, vibrate with an amplitude which is proportional to the impressed audio-frequency voltage. The latter is produced by

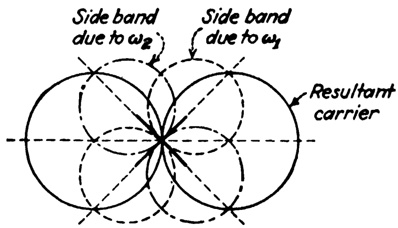


FIG. 299.—Radiation characteristics of double-modulation radio beacon when carrier currents in each loop antenna are in time phase.

the beats between the carrier and the side band and is therefore proportional to the product of these two. Multiplying the radius vectors of the side band and resultant carrier together in Fig. 299 gives the polar diagram of Fig. 300, the length of the radius vector being proportional to amplitude of vibration of the respective tuned reeds. It will be observed that only two courses are produced and that there is practically no signal received in the directions at right angles to these courses.

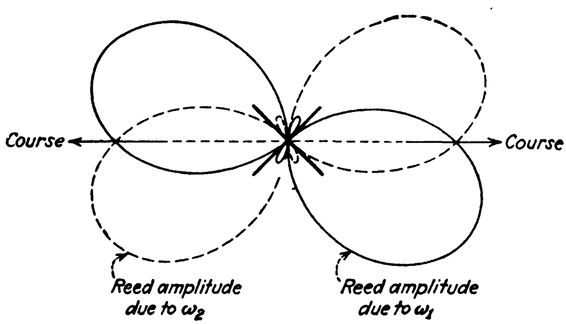


FIG. 300.—Amplitude characteristics of tuned reeds corresponding to radiation pattern of Fig. 299.

A four-course beacon can also be produced by adjusting the phase-control circuits of Fig. 298 so that the grid-excitation voltages impressed on the two modulated amplifiers are in quadrature. The carrier frequencies in the two antennas will therefore be 90

degrees out of time phase, and since they are also in space quadrature, a revolving field of carrier frequency will be produced in space which is of the same intensity in all directions. This is similar to the production of a revolving field in the case of a two-pole, two-phase induction motor, where the two pairs of poles are in space quadrature and the currents in the field coils are in time quadrature. The side bands of each loop, since the two modulating frequencies are different, will not combine and their space distribution remains the same as before. The radiation pattern in space is shown in Fig. 301 and the corresponding amplitudes of the tuned reeds in Fig. 302.

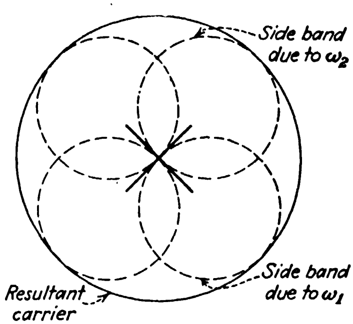


FIG. 301.—Radiation characteristics of double-modulation radio beacon when carrier currents are displaced 90 degrees in time phase.

In order to shift these courses from their 90-degree relationship so as to align them with existing airways, the current in one or the other of the two loops may be varied. The entire pattern in space may be rotated by rotating the secondary of the goniometer

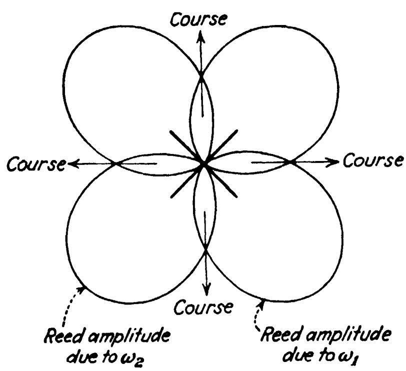


FIG. 302.—Amplitude characteristics of tuned reeds corresponding to radiation pattern of Fig. 301.

with respect to the primary. Another method of shifting the courses is to combine the radiation of a vertical antenna with that of the loops. The vertical antenna may be mounted along the central supporting mast and is excited with the same modu-

lated current as one of the loops.[17] The circular radiation pattern of the side bands from the vertical antenna will combine with the figure-eight pattern of this loop, and will reinforce one of the

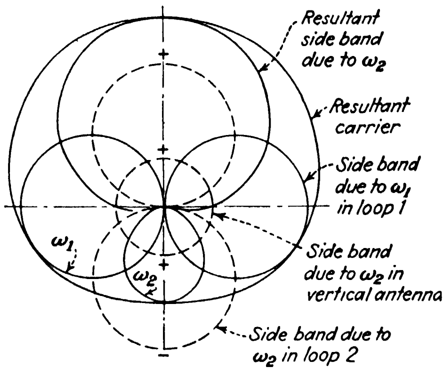


FIG. 303.—Radiation characteristics of crossed loop antennas combined with that of a vertical antenna.

lobes and tend to suppress the other, as shown in Fig. 303. The radiation pattern of the other loop will remain unchanged, since the frequency of its side bands differs from those of the vertical

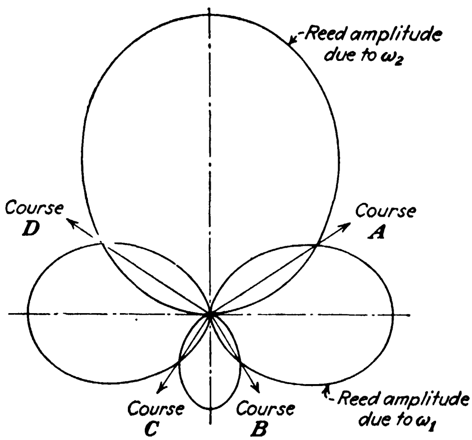


FIG. 304.—Amplitude characteristics of tuned reeds corresponding to radiation pattern of Fig. 303.

antenna. The corresponding amplitude characteristics of the tuned reeds are given in Fig. 304.

[17] H. DIAMOND, Applying the Visual Double-modulation Type Radio Range to the Airways, *Proc. I.R.E.*, vol. 17, p. 2158, December, 1929.

By the addition of a third modulating frequency at the transmitter, such as 108⅓ cycles, a 12-course beacon may be provided.[18] Three tuned reeds mechanically resonant to 65, 86⅔, and 108⅓ cycles are employed with the receiving set, only two of which are used on any one course. This series of frequencies results from the selection of 1300 r.p.m. as the driving speed of the modulation-frequency alternators which employ 6, 8, and 10 poles, respectively.

The chief advantages of the visual system are that it provides an optical comparison of the amplitudes of the received signals, and that the tuned reeds are relatively immune from disturbances caused by noise and interference. The transmitter is somewhat more complicated than the aural system and requires more careful maintenance and adjustment.

The foregoing discussion has all been based upon the use of loop antennas at the transmitter to secure the desired directional characteristics. It is also possible to use the circular radiation patterns around four vertical antennas to secure similar directional characteristics. These antennas are fed at the base by means of two-wire transmission lines coupled to the goniometer secondaries to obtain the proper phase relations. The transmission lines are nonradiating and there will therefore be no radiation due to currents flowing in the horizontal direction, as with the horizontal members of loop antennas. The elimination of radiation from horizontal conductors greatly reduces the sky wave and results in a considerable reduction in the apparent shifting of the beacon courses at night, when the deviations are always greatest.[19] The effect of the horizontally polarized sky wave in causing these random deviations has already been discussed in connection with direction finders.

180. Principles of Antenna Arrays.—For point-to-point communication at short wave lengths the desired horizontal directivity is usually obtained by an array of spaced antennas excited in the proper phase so as to obtain a preponderance of radiation

[18] H. DIAMOND and F. G. KEAR, A 12-course Radio Range for Guiding Aircraft with Tuned-reed Visual Indication, *Proc. I.R.E.*, vol. 18, p. 939, June, 1930; also W. E. JACKSON and S. L. BAILEY, *Proc. I.R.E.*, vol. 18, p. 2059, December, 1930.

[19] H. DIAMOND, On the Solution of the Problem of Night Effects with the Radio Range Beacon System, *Proc. I.R.E.*, vol. 21, p. 808, June, 1933.

in the desired direction. At longer wave lengths the required physical dimensions of the structure become prohibitively large and other methods of securing directivity are used. In the following discussion a transmitting antenna will be assumed, but it is to be understood that the directional characteristics obtained apply equally well to reception.

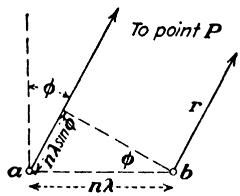


FIG. 305.—Notation for determining resultant field due to pair of spaced vertical antennas.

Consider the two similar vertical antennas at a and b in Fig. 305, separated a distance $n\lambda$ and both excited in time phase. The field at a point P, sufficiently remote from the two antennas so that lines drawn from a and b to P may be regarded as essentially parallel, will be the resultant of the individual fields due to a and b. The field due to the current in b will be

$$\mathcal{E}_b = \mathcal{E}_m \sin \omega\left(t - \frac{r}{c}\right) \tag{52}$$

while that due to a is

$$\mathcal{E}_a = \mathcal{E}_m \sin \omega\left(t - \frac{r + n\lambda \sin \phi}{c}\right) \tag{53}$$

where $\mathcal{E}_m$ is the maximum value of the field at a horizontal distance r from the antennas, and is given by either (27), (29), (30), or (31), depending on the current distribution along the vertical length of the antenna. It is assumed that the magnitudes of the two fields will be the same for equal currents in the antennas, which neglects the differences in the lengths of the two paths. The effect of this difference on the phase of the two fields at P cannot be neglected.

The resultant field at P will be

$$\begin{aligned} \mathcal{E}_r &= \mathcal{E}_b + \mathcal{E}_a \\ &= \mathcal{E}_m\left[\sin \omega\left(t - \frac{r}{c}\right) + \sin \omega\left(t - \frac{r + n\lambda \sin \phi}{c}\right)\right] \end{aligned} \tag{54}$$

From trigonometry

$$\sin A + \sin B = 2 \sin \tfrac{1}{2}(A + B) \cos \tfrac{1}{2}(A - B)$$

and (54) becomes

$$\mathcal{E}_r = 2\mathcal{E}_m \cos\left(\frac{\omega n\lambda \sin\phi}{2c}\right)\sin\left(\omega t - \frac{\omega r}{c} - \frac{\omega n\lambda\sin\phi}{2c}\right)$$
$$= 2\mathcal{E}_m \cos(\pi n \sin\phi)\sin\omega\left(t - \frac{r + \frac{1}{2}n\lambda\sin\phi}{c}\right) \qquad (55)$$

since $\omega/c = 2\pi/\lambda$. The maximum amplitude of the resultant field is

$$\mathcal{E}_r = 2\mathcal{E}_m \cos(\pi n \sin\phi) \qquad (56)$$

This resultant is the vector sum of the two fields due to a and b, as shown in Fig. 306. From the diagram the resultant is seen to be

$$\mathcal{E}_r = 2\mathcal{E}_m \cos\frac{\alpha}{2} \qquad (57)$$

where

$$\frac{\alpha}{2} = \pi n \sin\phi.$$

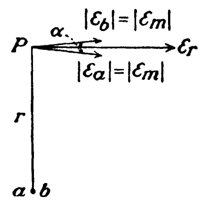


FIG. 306.—Resultant field due to antennas of Fig. 305.

Polar diagrams of the resultant field strength in the horizontal plane around two vertical antennas excited in phase are illustrated in Fig. 307 for various spacings. The equation of these curves is given by (56), the angle ϕ being measured from the vertical axis of the diagram in each case. With no separation between the two antennas the polar distribution curve is a circle. As the distance $n\lambda$ between them becomes larger, the field strength along the horizontal axis passing through the two antennas diminishes, becoming zero when the spacing becomes one-half wave length. As the spacing is increased still further, additional lobes appear.

If the currents in the two antennas are displaced from each other in time phase by an angle ψ, the angle α in Fig. 306 will be either increased or decreased, depending upon whether the current in antenna a leads or lags the current in b. The expression for the amplitude of the resultant field then becomes

$$\mathcal{E}_r = 2\mathcal{E}_m \cos\left(\pi n \sin\phi \pm \frac{\psi}{2}\right) \qquad (58)$$

the positive sign being used when the current in b *lags* the current in a. The negative sign is used when the current in b *leads* the

current in *a*. Polar diagrams for a time-phase displacement of $\psi = 90$ degrees (*b* lagging *a*) with two different spacings are shown in Fig. 308. A reversal of the time-phase relationship

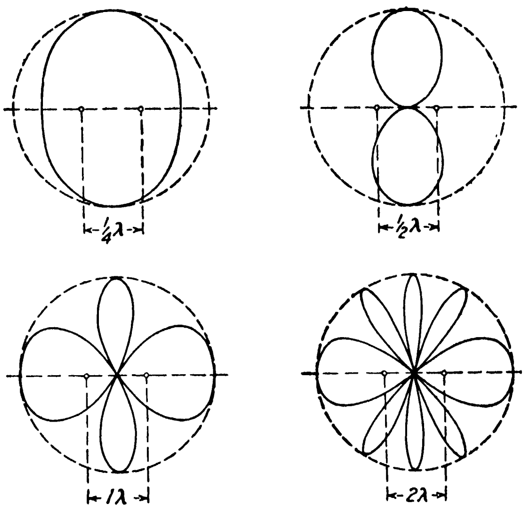


FIG. 307.—Field strength in horizontal plane around two vertical antennas separated by various distances and excited in phase.

reflects the diagram about the vertical axis, *i.e.*, the right and left sides are interchanged. A large number of these diagrams for various spacings and phase relations are given in the reference below.[20]

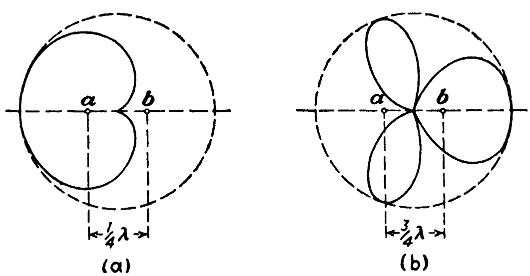


FIG. 308.—Field strength in horizontal plane around two vertical antennas when current in *a* leads current in *b* by 90 degrees.

The pattern of Fig. 308*a* can be obtained without actually energizing antenna *b*, for if antenna *a* is excited, a current will be

[20] R. M. FOSTER, Directive Diagrams of Antenna Arrays, *Bell System Tech. Jour.*, vol. 5, p. 292, April, 1926.

induced in b which will lag 90 degrees behind the current in a. In this way antenna b can be made to act as a reflector and confine the radiated energy to the desired direction. By using an array consisting of a number of spaced antennas backed by a similar row of reflectors, a sharply defined beam of radiation may be secured in a direction perpendicular to the line of the array. This is illustrated in Fig. 309 for an array consisting of 24 antennas spaced one-quarter wave length and backed by a similar reflecting structure having the same spacing.[21] The antennas are all fed in phase. Without the reflectors a bidirectional diagram would be obtained, each half of which would have had practically the

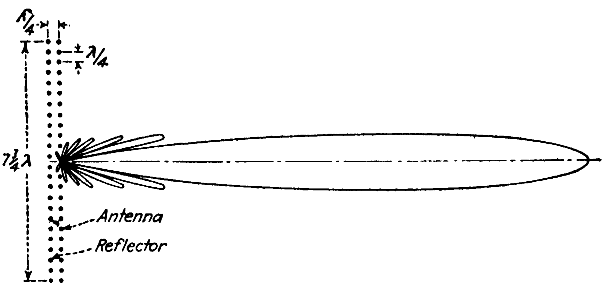


FIG. 309.—Directional diagram of array composed of twenty-four couplets spaced one-quarter wave length.

same general appearance and area of Fig. 309. The actual curve for an array when a quarter-wave reflector is used can always be obtained by multiplying the polar curve of the array alone by the cardioid curve of Fig. 308a for a single antenna and a single reflector wire one-quarter wave length behind it.

The directional characteristic of an array is but little affected by the individual spacing $n\lambda$ between wires when this distance becomes less than $\lambda/2$, assuming the phase of all the currents to be the same. As the total length of the array is increased, the directive properties become more marked and the radiation is confined to an increasingly narrow beam, as illustrated in Fig. 310. Curve a shows a portion of the polar diagram for an array having a total length of 2λ, while b is for an array length of 10λ. Only the right-hand portions of the diagrams are shown.

A receiving antenna which is too highly directive is unsatisfactory because of random deviations in the direction of arrival of

[21] G. C. SOUTHWORTH, Certain Factors Affecting the Gain of Directive Antennas, *Proc. I.R.E.*, vol. 18, p. 1502, September, 1930.

the sky wave, previously mentioned in connection with direction finders. Thus a shift of only a few degrees in the bearing of the received wave would cause a considerable reduction in signal strength, even though the field strength of the wave remained constant.

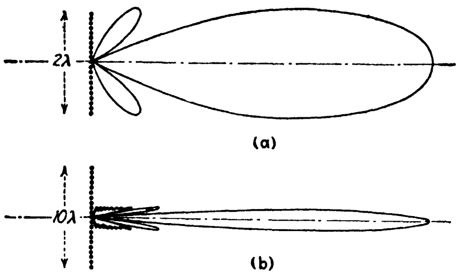


FIG. 310.—Effect of array-length upon horizontal directive characteristics. Only the right-hand portion of the polar diagram is shown.

181. Determination of the Polar Diagram for an Array.—Consider an array of N vertical antennas, each separated by a distance $n\lambda$, as shown in Fig. 311*a*, and all excited in phase. At a point P remote from the array the resultant field $\mathcal{E}_R$ will be the vector sum of the individual fields $\mathcal{E}_m$ produced by each antenna. Each of

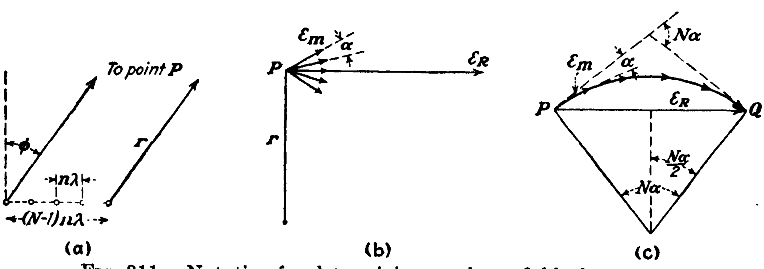


FIG. 311.—Notation for determining resultant field of an array.

these individual vectors will be displaced from its neighbors by an angle α which is equal to, from (57)

$$\alpha = 2\pi n \sin \phi \tag{59}$$

The geometrical relations are more readily seen in (*c*). If the distance $n\lambda$ is small, these vectors approximate the arc of a circle, the chord PQ being equal to $\mathcal{E}_R$. The angle subtended by the arc at the center of the circle is equal to the angle of phase differ-

ence between the first and last vectors and is equal to $N\alpha$. If the radius of the circle is a, the length of the arc PQ will be the angle $N\alpha$ in radians multiplied by the radius a or

$$\text{arc } PQ = aN\alpha$$

while

$$\text{chord } PQ = 2a \sin \frac{N\alpha}{2}$$

The ratio of these two lengths is

$$\frac{\text{chord } PQ}{\text{arc } PQ} = \frac{\sin \dfrac{N\alpha}{2}}{N\alpha/2} \tag{60}$$

Since $\mathcal{E}_R$ is less than $N\mathcal{E}_m$ by this factor,

$$\begin{aligned} \mathcal{E}_R &= N\mathcal{E}_m \frac{\sin \dfrac{N\alpha}{2}}{N\alpha/2} = \mathcal{E}_m \frac{\sin \dfrac{N\alpha}{2}}{\alpha/2} \\ &= \mathcal{E}_m \frac{\sin (N\pi n \sin \phi)}{\pi n \sin \phi} \end{aligned} \tag{61}$$

where the fraction of a wave length n separating the individual antennas is small.

The diagram of Fig. 311c will be recognized as being identical with that used in the determination of the breadth factor of an alternator having a smoothly distributed winding on the armature. In the case of the alternator, $\mathcal{E}_m$ would represent the voltage induced in a coil and $N\alpha$ would be the width of the phase belt measured in electrical degrees; N being the number of coils and α the angle between them. The expression for the breadth factor in this case is identical with (60). When the coils of the alternator are distributed in N slots spaced α degrees apart, the breadth factor becomes[22]

$$\frac{\sin \dfrac{N\alpha}{2}}{N \sin \dfrac{\alpha}{2}} \tag{62}$$

[22] R. R. LAWRENCE, "Principles of Alternating Current Machinery," p. 41.

This is evidently similar to the case where the spacing $n\lambda$ between antennas is too great to consider the individual vectors $\mathcal{E}_m$ of Fig. 311*c* as equivalent to the arc of a circle, and they must therefore be treated as a portion of a polygon. Substituting the more accurate expression of (62) in place of (60), (61) becomes

$$\mathcal{E}_R = \mathcal{E}_m \frac{\sin (N\pi n \sin \phi)}{\sin (\pi n \sin \phi)} \tag{63}$$

which is correct for any value of n. As n becomes small, the angle may be taken for the sine in the denominator of (63), giving the expression obtained in (61).

In plotting polar distribution curves like those of Figs. 309 and 310, we are usually concerned with only the relative magnitudes of $\mathcal{E}_R$ in the various directions. If the maximum length of the vector $\mathcal{E}_R$ is taken as unity, the length of the radius vector for various values of ϕ in (63) is

$$\rho = \frac{\sin (N\pi n \sin \phi)}{N \sin (\pi n \sin \phi)} \tag{64}$$

The angular position of the various minima can be ascertained by considering the various values of ϕ which will make the numerator of (64) equal to zero. The first minimum will occur when $N\pi n \sin \phi = \pi$ or when

$$\sin \phi = \frac{1}{Nn} \tag{65}$$

Therefore an array composed of 8 antennas spaced a quarter of a wave length—an array length of approximately 2λ—the position of the first minimum will be given by $\sin \phi = \dfrac{1}{8 \times 0.25} = 0.5$, or $\phi = 30$ degrees. The second minimum will occur when $N\pi \sin \phi = 2\pi$, etc. As the quantity Nn in (64) becomes larger, the number of minima increases and the angular separation between them becomes less. Since each lobe is bounded by a pair of minima, the number and angular widths of the major and minor lobes may be readily determined.

When a parasitic antenna is used as a reflector, as in Fig. 309, the resultant characteristic can be obtained by multiplying the polar characteristic of the array alone by the characteristic of a

system composed of a single antenna and a single reflector wire a quarter wave length behind it. Such a characteristic is given in Fig. 308*a*, the equation of the curve being, from (58)

$$\rho = \cos\left[\frac{\pi}{4}(\cos\ \phi - 1)\right] \tag{66}$$

The equation of an array backed by a reflector spaced $\lambda/4$ behind the antennas will therefore be

$$\rho = \frac{\sin\ (N\pi n \sin\ \phi)}{N \sin\ (\pi n \sin\ \phi)} \cos\left[\frac{\pi}{4}(\cos\ \phi - 1)\right] \tag{67}$$

where N = number of individual antennas (not including reflectors).
n = spacing between antennas in wave lengths.
ϕ = angle between radius vector ρ and reference axis.

In plotting polar diagrams from (67), negative values of ρ have no physical significance and are to be plotted in a positive sense. Expressions for more complicated arrays, together with their polar diagrams, are given in Southworth's paper.[21]

The theoretical spacing of the reflector a quarter wave length behind the row of antennas must be modified slightly in practice because of finite conductivities and also because of the reactions of the two structures upon each other. If the separation between the row of antennas and reflectors is exactly $\lambda/4$, the reflectors must be slightly detuned from resonance so that the parasitic currents induced will be of exactly the proper phase to reinforce the radiation in the desired direction.[23]

When the parasitic reflecting antenna is relatively close to the transmitting antenna the induction field will induce an appreciable voltage in the former which will be 90 degrees out of phase with that produced by the radiation field. The phase of the resultant induced current can be altered by making the reactance of the reflector slightly inductive or capacitive by altering its length.

[21] *Loc. cit.*

[23] R. M. WILMOTTE and J. S. MCPETRIE, A Theoretical Investigation of the Phase Relations in Beam Systems, *Jour. I.E.E.* (London), vol. 66, p. 949, September, 1928.

182. Types of Arrays.—The type of structure just considered is known as the broadside array. If this row of vertical antennas is excited so that there is a progressive phase difference between the currents of the adjacent antennas, the radiation will be concentrated along the axis of the array instead of at right angles to it. This type of structure is known as an end-fire array. The time-phase difference between the currents in the adjacent antennas is usually made equal to their space-phase separation. Thus a separation of $\lambda/4$ would require a progressive phase shift of 90 degrees in the excitation currents along the array. The radiation is concentrated in a unidirectional beam coinciding with the line of the array and is directed toward the end in which the phase lags when the spacing between the elements is an odd number of quarter wave lengths and the phasing is an odd number of quarter periods. When the spacing and phasing are both an even number of quarter wave lengths and quarter periods, a bidirectional characteristic is obtained.

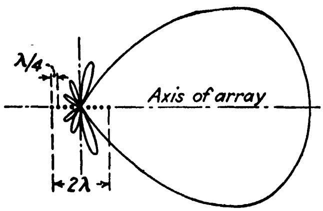


FIG. 312.—Directional diagram of end-fire array having a total length of 2λ.

A typical directional characteristic in the horizontal plane for an end-fire array having a total length of 2λ is shown in Fig. 312. The beam is much broader than that for a broadside array of the same total length, as will be seen by comparing this diagram with that of Fig. 310*a*. The directivity improves as the total length of the array is increased, but at a much slower rate than with the broadside array. The directivity in the vertical plane is similar to the horizontal pattern.

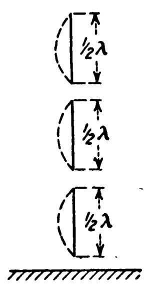


FIG. 313.—End view of arrays stacked one above the other and excited in phase.

A considerable improvement in vertical directivity results if a series of arrays are stacked one above the other and all excited in the same phase, as shown by an end view in Fig. 313. This avoids the interference between the fields in a horizontal direction produced by the alternate half waves of opposite sign, which is characteristic of an ordinary harmonic antenna several waves in length. By doing away with adjacent half waves which are in

phase opposition, the high-angle radiation is greatly reduced and the energy is concentrated in a horizontal beam. A unidirectional characteristic may be secured by a similar tiered group of reflectors, placed a quarter wave length behind the array of Fig. 313.

Various combinations of these types are used in practice. Figure 314 shows the type of vertically polarized transmitting array that has been used for the short-wave channels of the transoceanic telephone station at Lawrenceville, N. J. Only one of the several bays composing the structure is shown, and the reflector curtain of similar construction located a quarter wave length behind the antenna curtain has been omitted for clarity.

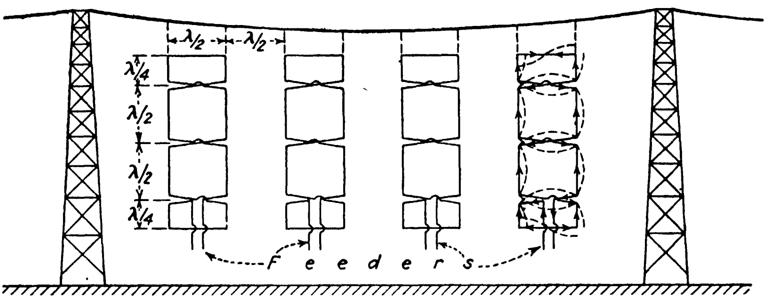


FIG. 314.—Transmitting array used for short-wave transoceanic telephone.

The conductors are so arranged that the standing half waves on the vertical members are all in phase, similar to the stacked antennas of Fig. 313. The fields due to the alternate half waves in the horizontal members, which are of opposite phase, cancel each other owing to their close proximity. The feeders supply current to each unit and are terminated and adjusted as to length so that all the units are excited in exactly the same time phase. The closed metallic circuits of each unit enables 60-cycle heating current to be sent through the conductors to remove deposits of sleet.[24] The formation of ice on the conductors during severe sleet storms may produce an appreciable detuning effect as well as dangerous mechanical stresses.

A somewhat similar array developed by T. Walmsley[25] and used by the British Post Office is shown in Fig. 315*a*. It consists of two

[24] E. J. STERBA, Theoretical and Practical Aspects of Directional Transmitting Systems, *Proc. I.R.E.*, vol. 19, p. 1184, July, 1931.

[25] Beam Arrays and Transmission Lines, *Jour. I.E.E.* (London), vol. 69, p. 299, February, 1931.

parallel rows of vertically polarized radiators spaced one-half wave length apart. The currents in the vertical members of each row are in phase with each other, but there is a phase difference of 180 degrees between the two rows which will result in a bidirectional characteristic if no reflector is used. This reflector consists of a curtain of half-wave antennas arranged in two vertical stacks placed a quarter wave length behind the second row of antennas. The fields due to the horizontal elements cancel each other owing to their proximity.

The array in Fig. 315*b* is composed of a curtain of horizontal radiating elements, as the fields produced by the vertical elements cancel in this case. The resultant radiation is now horizontally

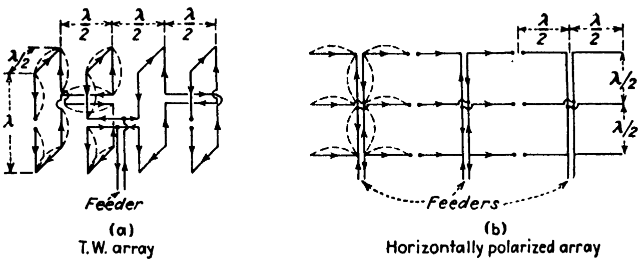


FIG. 315.—Other forms of antenna arrays. The reflectors are omitted for simplicity.

polarized. There seems to be little difference in the results obtained between vertically and horizontally polarized radiation for short-wave use, since reception at great distances is possible only by means of the sky wave which always experiences more or less shifting of the plane of polarization. A reflecting curtain is required by this type of antenna if a unidirectional characteristic is desired.

183. Arrays for Reception.—While any of the preceding structures can also be used for reception, receiving arrays are usually somewhat simpler in construction. They are also not so tall since the stacking of several tiers of elements one above the other, which is characteristic of most transmitting arrays, is not necessary for reception. The increased height, owing to stacking, is for the purpose of increasing the vertical directivity so as to concentrate the radiated energy in the form of a beam directed toward the horizon. But the direction of arrival of the waves at the receiver varies and may be inclined considerably above the

horizontal at times, so that a highly directive characteristic in the vertical plane would cause poor reception on such occasions.

A widely used receiving array of the broadside type is illustrated in Fig. 316. Maximum response results when the received waves are normal to the plane of the array. The amount of horizontal directivity will depend upon the total length of the structure, as illustrated in Fig. 310. This should not be made too great because of variations in the bearing of the received wave.

A wave traveling in the direction from the antenna to the reflector will induce an e.m.f. in the latter which lags 90 degrees behind the e.m.f. induced in the antenna. Since the feeder line from the reflector is a quarter wave length longer than that from the antenna, the two e.m.fs. will be additive between the feeder terminals A and R. But a wave arriving from the rear will

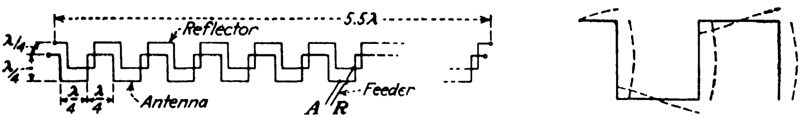


FIG. 316.—Broadside receiving array.

induce an e.m.f. in the reflector 90 degrees ahead in phase of that induced in the antenna, and the resultant voltage between A and R will now be zero.

An array of these dimensions will deliver almost forty times as much power to the receiver as that delivered by a single vertical half-wave antenna. If static were distributed uniformly around the array, a similar improvement in the noise-to-signal power ratio would also be observed. An even greater ratio can be obtained in the case of directional static if one of the minima in the directional characteristic can be pointed toward the origin of the disturbances. Short-wave static is highly directional on many occasions.

184. Practical Aspects of Antenna Arrays.—One very important advantage obtained by the use of a directional transmitting antenna is in the greatly increased field strength produced at the receiver for a given amount of power furnished to the antenna. The amount of gain obtained by the use of an array can be conveniently expressed as the ratio of the power required by the comparison antenna, to the power required by the array when equal field strengths are produced in the chosen direction. The antenna used for comparison purposes may be either a simple

vertical half-wave doublet, or else one of the array elements. The gains obtainable are functions of element spacing and array length, and have been investigated by Sterba[24] for a number of cases. Typical results are shown in Fig. 317 for a broadside array of vertical elements compared to a single element. Thus an array 14λ in length will have a gain of almost 16 db over a simple vertical antenna, which means that the latter would have to be excited with almost fifty times the power required by the

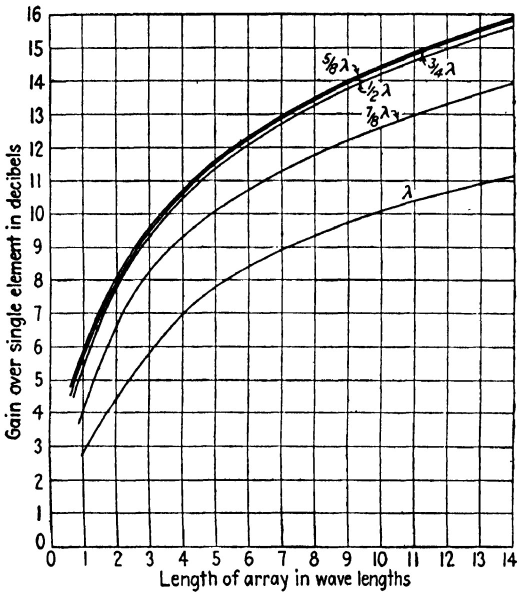


FIG. 317.—Gain in decibels of vertical broadside array over single element as a function of array length, for various element spacings.

array to produce the same field strength at the receiver. The power ratio is almost a linear function of the array length. However, the economic considerations involved in the cost of a large array with its numerous supporting towers as compared with the cost of producing the desired field strength by an increased amount of power furnished to a smaller array must be carefully investigated.

One serious disadvantage possessed by these arrays is that the operating frequency is intimately related to the lengths and spacings of the elements and cannot, therefore, be adjusted through a wide range of frequencies to meet various operating

[24] *Loc. cit.*

conditions. For long-distance communication in daylight the optimum wave length is considerably shorter than at night, and it also varies between summer and winter. Consequently, several wave lengths must be utilized if reliable 24-hr. communication is to be maintained throughout the year. This calls for a separate array for each wave length used. Three arrays are used for the transatlantic telephone at Lawrenceville (Fig. 314) which are supported by a row of 19 towers spaced 250 ft. apart and 180 ft. high, in a line at right angles to the direction of England.

185. Directional Antennas Employing Long Wires.—The radiation patterns produced by wires several waves in length

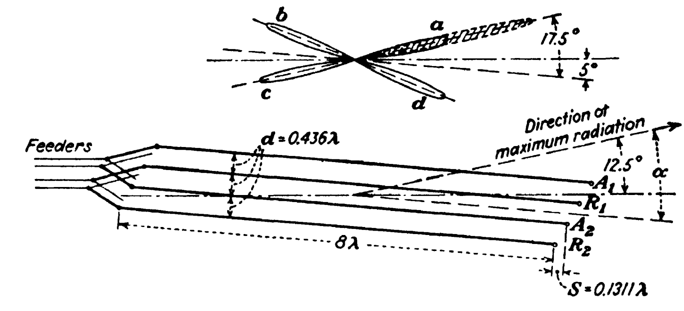


FIG. 318.—Directive antenna using long tilted wires.

(Fig. 287) may be combined to produce a directional characteristic comparable to those of the rectangular arrays just discussed.[6] A wire 8λ in length produces four major lobes of radiation making an angle of 17.5 degrees with the wire, as shown in the upper portion of Fig. 318. The minor lobes are omitted for simplicity. If the wire A_1 is inclined with respect to the earth's surface, the axis of lobe a may be given any inclination desired with respect to the horizontal, in this case, 12.5 degrees. A second wire A_2, parallel to the first and in the same vertical plane, if excited by a current in phase opposition to that in A_1, will annul the field produced by A_1 at all points in a plane passing through R_1 and normal to the plane containing the two wires. This is evident, since any point in this plane is equidistant from each wire and the two currents in the wires are 180 degrees out of phase. A portion of the desired horizontal directivity has thus been secured.

[6] *Loc. cit.*, p. 428.

The two lobes b and d may be removed by axially displacing A_2 by an amount $2s$, so that in the direction along the axis of b and d, any point on one wire is in phase opposition to the corresponding point on the other. The resultant field will then be zero along this line and each of the two major lobes will be split into a small pair of minor lobes. In order to remove the remaining lobe c, a pair of reflector wires R_1 and R_2 are inserted below A_1 and A_2, as shown. These are also excited with currents that are in phase opposition to each other, but which lead those in A_1 and A_2 by

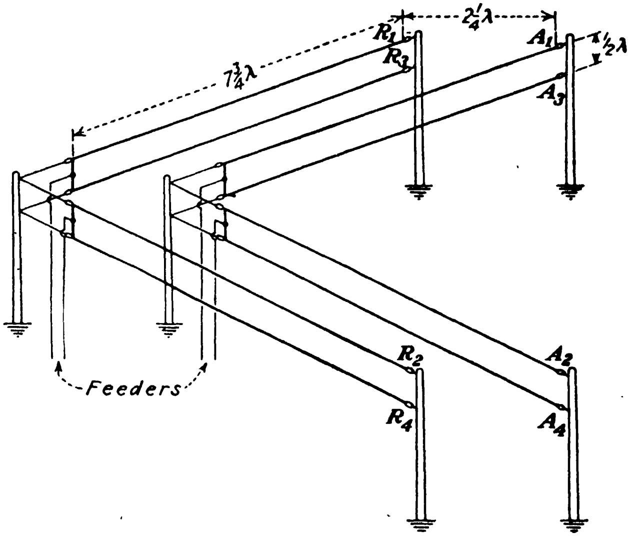


FIG. 319.—Directive V-type of long wire antenna.

90 degrees. This reinforces lobe a, as represented by the shaded area, and subdivides c into a pair of minor lobes. The action of these reflectors is similar in effect to that of Fig. 308a. A unidirectional characteristic is thus obtained. The vertical and horizontal radiation patterns are similar, which is a characteristic of end-fire arrays.

The required endwise displacement s is given by

$$s = \frac{\lambda}{4} \frac{\sin \alpha}{\sin 2\alpha} \tag{68}$$

and the distance between wires by

$$d = \frac{\lambda}{4} \frac{\cos \alpha}{\sin 2\alpha} \tag{69}$$

where α is the angle between the axis of maximum radiation and the wire axis; 17.5 degrees in the case of a wire 8λ in length.

Another type of antenna using long-wire radiators which is partially folded upon itself from the center is shown in Fig. 319. The currents in A_1 and A_2 are in phase opposition, causing the radiation, which will be bidirectional, to be concentrated along a line bisecting the "V." A second pair of wires, A_3 and A_4, located approximately a half wave length below and excited in phase with the first pair, serve to reduce the high-angle radiation. In order to secure a unidirectional beam, two pairs of reflector wires are spaced an odd number of quarter wave lengths (usually $2\frac{1}{4}\lambda$) behind the first set, and are excited with currents in quadrature to those in the front antennas. These currents may be either leading or lagging, depending upon whether the desired direction of transmission is from R to A, or A to R. Increased horizontal directivity may be had by joining two V sections to form a W.

186. The Beverage or Wave Antenna.[26]—This type of antenna consists of a long horizontal wire varying from a half to several waves in length and is used for directional reception at long waves. In its simplest form it consists of a conductor several miles long (depending on the wave length of the signals to be received) and is pointed in the direction of the transmitting station. The height is comparatively low, the usual construction being to mount the conductor on ordinary telephone poles. The far end is grounded through an impedance Z_L equal to the characteristic or surge impedance of the structure, considering it to be a one-wire transmission line with a ground return. This terminating impedance is approximately a pure resistance having a value of $\sqrt{L/C}$ ohms, where L and C are the inductance and capacitance of the conductor per unit length.*

[26] H. H. BEVERAGE, C. W. RICE, and E. W. KELLOGG, The Wave Antenna, *Trans. A.I.E.E.*, vol. 42, p. 215, February, 1923.

* The accurate expression for the characteristic impedance is

$$Z_L = \sqrt{\frac{R + j\omega L}{G + j\omega C}},$$

which reduces to the above expression if R and G per unit length are negligible in comparison with ωL and ωC. See Sec. 194.

A wave traveling in the direction shown in Fig. 320 will induce a small e.m.f. in the end of the conductor adjacent to Z_L. This impulse travels along the wire, and if its velocity is practically the same as that of the electromagnetic wave in space, the voltage wave will continue to build up as it travels toward the receiving end, as it is being continuously supplied with energy in the proper phase from the wave traveling in space. This process is analogous to the building up of water waves in the direction of the wind. If the velocity of travel in the conductor is somewhat slower than the space wave, interference effects develop, the wave on the wire building up for a certain distance and then decreasing in amplitude. The voltage induced in the wire is due to the fact

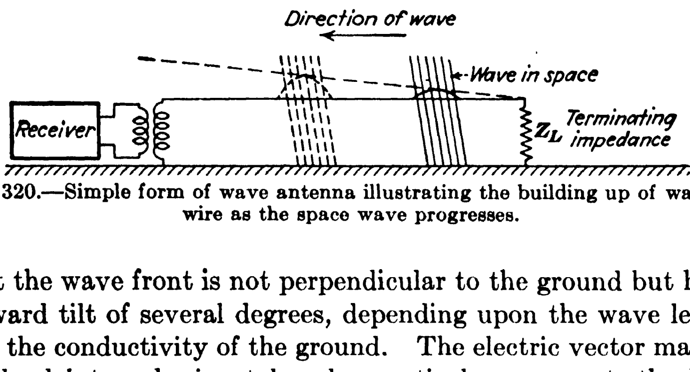


FIG. 320.—Simple form of wave antenna illustrating the building up of wave on wire as the space wave progresses.

that the wave front is not perpendicular to the ground but has a forward tilt of several degrees, depending upon the wave length and the conductivity of the ground. The electric vector may be resolved into a horizontal and a vertical component, the horizontal component being the one which is responsible for the wave induced in the wire.

A wave traveling along the wire in the opposite direction will likewise produce a traveling wave of increasing magnitude on the wire, but it will be entirely absorbed without reflection by the terminating impedance, as no reflections can occur at the end of a transmission line terminated in an impedance equal to the characteristic impedance of the line. If the line were merely grounded, a reflection would occur which would travel back to the receiver, and a bidirectional characteristic would be obtained. Practically no voltage will be induced by a wave traveling at right angles to the antenna because of the short vertical height. The discrimination against waves at right angles becomes less at shorter wave lengths since the vertical height becomes more comparable to the horizontal length.

In order to have the terminating impedance Z_L located at the same end of the antenna as the receiving set for convenience of adjustments, the circuit of Fig. 321 is used. The two wires act in parallel as an antenna and also act as a transmission line to return the signal to the receiver. This is accomplished by means of the reflecting transformer, connected as shown. The induced current in the antenna flows in opposite directions in the two halves of the transformer primary and through the secondary coils to ground. But the current in these secondary coils will induce voltages in the two halves of the former primary, which act in the same direction in both coils, as shown by the dotted arrow. The two antenna conductors will then act as an ordinary transmission line and transmit this voltage to the receiver. The

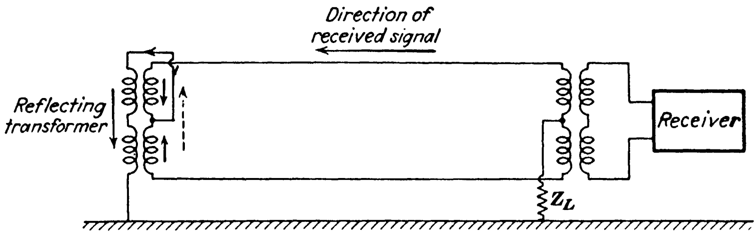


FIG. 321.—Wave antenna arrangement for locating the receiving set at the same end as the terminating impedance.

voltage induced in the two antenna wires by a wave traveling in the opposite direction will be entirely absorbed at the receiver end by Z_L and no reflections will take place.

The directional pattern of a wave antenna becomes sharper as its length is increased. The length is usually limited at very long wave lengths by physical considerations. Furthermore, as the length is increased beyond the optimum value, the output voltage begins to diminish owing to the increased lag in phase of the wave on the wire behind the wave in space. In some cases lengths greater than optimum must be used to secure the desired directivity. The horizontal directivity is not so good as that of the preceding types of arrays of the same length measured in waves. An array of several wave antennas is used to secure the desired directivity at the receiving station for the long-wave channel of the transatlantic telephone at Houlton, Maine.[27]

[27] A. BAILEY, S. W. DEAN, and W. T. WINTRINGHAM, The Receiving System for Long-wave Transatlantic Radio Telephony, *Proc. I.R.E.*, vol. 16, p. 1645, December. 1928.

induced in the successive elements, causing the resultant currents in R to add up to the same maximum resultant as in Fig. 322*a*. For any angle of tilt ϕ there will be a wire length which will make this resultant a maximum. This occurs when the angle ϕ is such that the length of wire is $\lambda/2$ longer than its horizontal projection, assuming the direction of wave propagation to be horizontal.

The optimum tilt angle is plotted in Fig. 324 as a function of the wire length. It will be observed that as the length increases, the optimum value of ϕ changes slowly. For example, an antenna designed for a frequency such that the wire was 8 waves long is

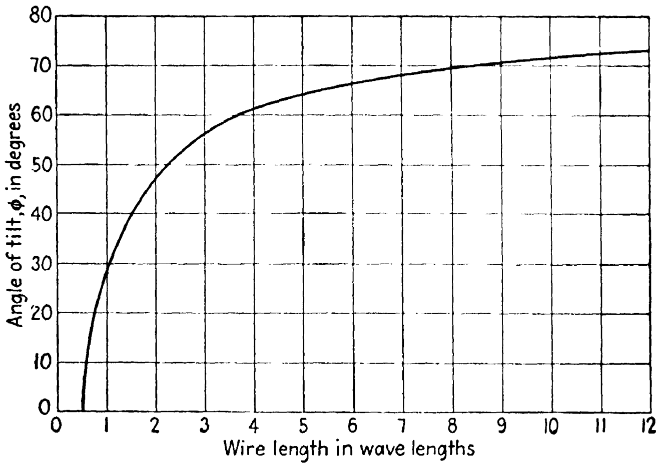


FIG. 324.—Optimum tilt angle of tilted-wire antennas.

to be used on a higher frequency where the antenna length is now 12 waves. Under this condition, the tilt will be only a few degrees less than optimum, and the reduction in the received current would be inappreciable. This enables a tilted wire to be used on a fairly wide range of wave lengths with equally good results. As ϕ approaches 90 degrees, the performance of the tilted wire approaches that of the wave antenna.

If two tilted antennas are arranged to form an inverted V, as in Fig. 325, a unidirectional characteristic can be obtained. The end remote from the receiver R is grounded through a resistance equal to the characteristic impedance to avoid reflections. A wave coming from one direction will produce maximum current in R, while a wave from the opposite direction will have a result-

ant equal to zero, as shown, resulting in a unidirectional characteristic having an infinite front to back ratio. It can be shown that this ratio exists whenever the length of each tilted wire is an odd integral multiple, greater than unity, of $\lambda/4$, provided ϕ has its optimum value. This favorable condition is attained at only one frequency, and at other values of frequency the ratio is finite. However, by adjusting the terminating impedance at the far end so as to permit an equal and opposite signal to be reflected from this point, complete cancellation can again be obtained.

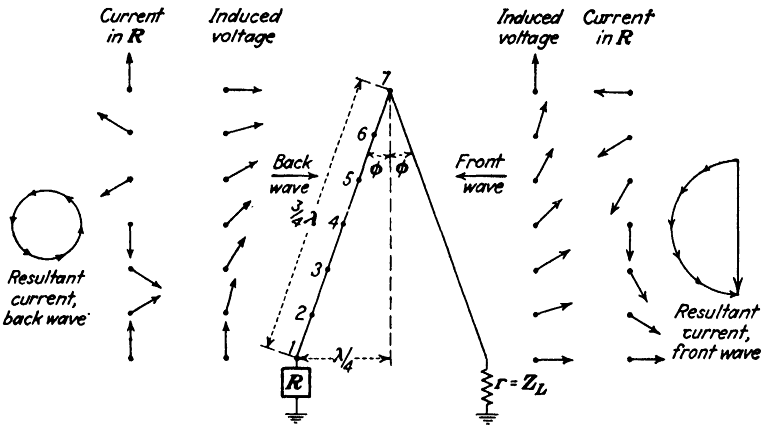


FIG. 325.—Resultant current in an inverted-V antenna. The phase of the elementary currents produced by the back wave changes twice as rapidly as that produced by the front wave, and cancels in R.

In practice the terminating impedance is usually adjusted to some compromise value giving a high front to back ratio over the entire useful range of frequencies.

188. Horizontal Rhombic or Diamond-shaped Antenna.[29]—This type of antenna, devised by Bruce, consists of two horizontal V antennas arranged as shown in Fig. 326. No ground connections are required in this case, which does away with the variations in ground-contact resistance caused by variable weather conditions. In Fig. 325 the contact resistances of the grounds were in series with the terminating resistances at each end, so that any changes in ground resistance caused variations in the total value of terminating impedance.

[29] E. BRUCE, A. C. BECK, and L. R. LOWRY, Horizontal Rhombic Antennas, *Proc. I.R.E.*, vol. 23, p. 24, January, 1935.

In practice the two sides of the diamond are composed of a pair of conductors in parallel having a variable spacing between them, as shown in the elevation view of Fig. 326. This construction serves to offset the variation in the distributed inductance and capacitance per unit length of conductor that otherwise would be caused by the variable spacing between the two opposite sides of the diamond. Without this compensation the characteristic impedance would be different at various points along the antenna.

The operation of this antenna is exactly the same in principle as the tilted-wire structures discussed in the preceding section, except that the horizontal component of the received wave is

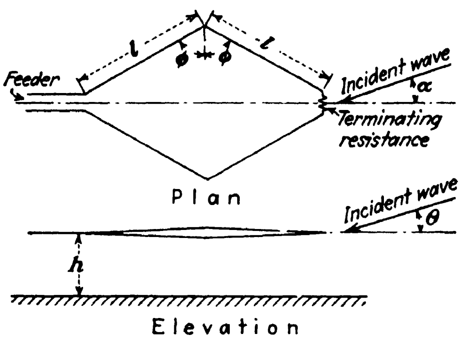


FIG. 326.—Horizontal rhombic or diamond-shaped antenna.

responsible for the production of the signal. With short waves the horizontal component is often stronger than the vertical component. These antennas are suitable for transmitting, and when so used produce horizontally polarized waves. The terminating resistance in this case is called upon to dissipate an amount of energy equal to that which would have been radiated in the backward direction. A long two-wire iron transmission line, shorted at the far end, is usually employed for this purpose.

The directional characteristics of a typical antenna of this type are shown in Fig. 327. Figures *a*, *b*, and *c* show the effect of frequency variations on the horizontal directivity, the antenna dimensions remaining the same in all cases. The variations in vertical directivity with frequency will be of about the same order of magnitude. It will be observed that these patterns compare very favorably with those of the more complicated arrays previously discussed. The structure is also much simpler and cheaper

to construct, as the height above the ground will be only from 40 to 80 ft., depending upon the frequencies to be transmitted or received. This feature, combined with a useful frequency range of about 2½:1, enables a single rhombic antenna of fixed dimensions to be used for both the day and night frequencies. This is in marked contrast to the older types of directive antennas which required a separate array for each wave length used, the cost of each array being considerably greater than that of the single rhombic type.

Like other horizontal antennas for short-wave reception, this type is somewhat more immune from local noise caused by electrical apparatus than those designed to receive vertically

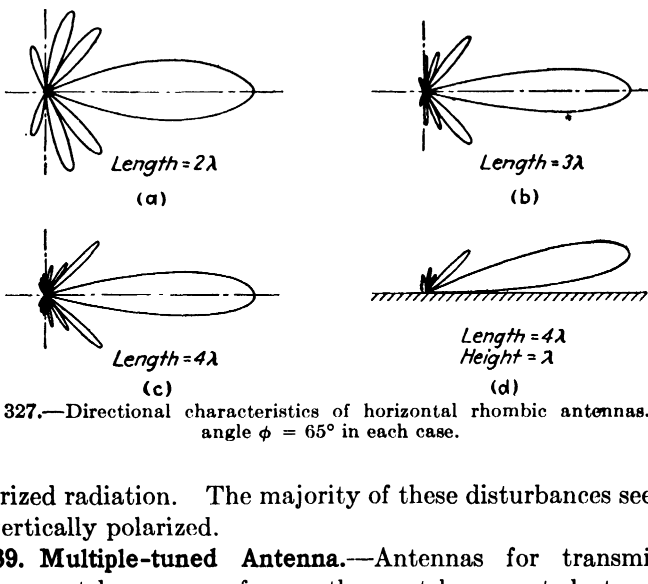


FIG. 327.—Directional characteristics of horizontal rhombic antennas. Tilt angle ϕ = 65° in each case.

polarized radiation. The majority of these disturbances seem to be vertically polarized.

189. Multiple-tuned Antenna.—Antennas for transmitting purposes at long waves frequently must be operated at a wave length considerably above the natural wave length of the structure, because of the practical considerations which limit the maximum physical size of the antenna. When so operated, the radiation resistance is rather low, and if a reasonable value of radiation efficiency is to be secured, the other items of resistance must be kept as low as possible. When the station is located where the conductivity of the soil is rather poor, the resistance of the ground connection is apt to be the major portion of the total

antenna resistance. Under these conditions the multiple-tuned antenna[30] may be used to advantage.

The construction is essentially that of an inverted-L antenna with the addition of a number of tuned vertical down leads attached to the flat-top portion and spaced at regular intervals, as illustrated in Fig. 328. These vertical leads are provided with individual tuning inductances at the base which are adjusted so that the currents in the vertical portions are all in phase. The effect is that of a number of vertical antennas connected in parallel and excited in phase. Since each vertical lead is provided with an individual ground, the total ground resistance will be that of the individual grounds all connected in parallel. The antenna structure is similar to a vertical broadside array except that the horizontal length measured in wave lengths is too short to produce any appreciable directive properties. There will also be radiation from the flat-top portion. By tuning the vertical members so as to produce a progressive phase displacement in the current in each lead, the properties of a very short end-fire array could be secured, but with reduced radiation resistance of the antenna as a whole. Increased radiation efficiency is the item ordinarily sought for in this type of construction and the directional possibilities are not ordinarily utilized.

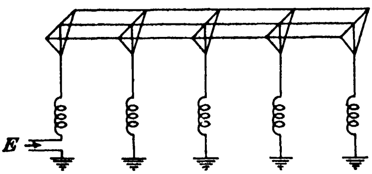


Fig. 328.—Multiple-tuned antenna.

190. Radio-frequency Transmission Lines.[31]—It is usually undesirable in transmitting stations to have the transmitting equipment housed immediately beneath the antenna. Not only would the presence of the building produce absorption losses in the radiated energy, but there would also be possibilities of feed-back to the transmitting apparatus. Metal work within the building such as railings, supports, exposed wiring, and the like, would have radio-frequency voltages induced in them which might be annoyingly large unless these items were properly grounded or shielded. For these reasons it is usually the practice

[30] E. F. W. ALEXANDERSON, Transatlantic Radio Communication, *Trans. A.I.E.E.*, vol. 38, p. 1089, Part II, 1919.

[31] E. J. STERBA and C. B. FELDMAN, Transmission Lines for Short-wave Radio Systems, *Proc. I.R.E.*, vol. 20, p. 1163, July, 1932.

to locate the transmitting apparatus some distance away from the antenna and energize the latter by means of a suitable transmission line. This enables several transmitters to be housed in the same building, each being connected by a transmission line to its individual antenna. Most of the directional transmitting arrays previously described must be excited at various points by currents of the proper phase, which requires a number of properly terminated feeder lines. The losses in a well-designed transmission line are small and transmission distances of several thousand feet may be used if necessary.

The transmission line may be either resonant or nonresonant and is constructed so as to be nonradiating. This latter feature is achieved by the use either of two parallel wires placed fairly close together and balanced with respect to capacitances to ground, or of concentric tubular conductors. The open-wire line is less expensive and relatively simple to construct, but it is not so satisfactory from the standpoint of extraneous e.m.fs. induced in the circuit as the coaxial conductor, particularly in transmission lines for highly directive receiving antennas. Here, extraneous pick-up on the part of the line to the receiving unit may destroy most of the advantages of the directive antenna, particularly as to the ratio of signal to noise. The possibilities of cross talk between adjacent lines is also greater with the open-wire construction. If the outer sheath of the coaxial line can be grounded at frequent intervals complete shielding is obtained. In some cases lines of this type are buried in the ground.

191. Resonant Transmission Lines.—For short waves where a simple Hertz antenna is used in close proximity to the transmitter a resonant transmission line is usually employed, typical examples of which are illustrated in Fig. 329. The current-feed arrangement shown in (*a*) uses a line approximately $\lambda/2$ in length, or some integral multiple of this value. The current and voltage distributions are as shown, the current being a maximum at the transmitting end of the line. The impedance looking into this end of the line, neglecting the line losses, will be exactly equal to the terminating impedance offered by the antenna, which in the case of a horizontal resonant half-wave antenna will be theoretically its radiation resistance of 74 ohms. The actual value will differ from this because of ohmic and absorption losses, and also because of reflection from the ground. The effect of the

latter will depend upon the height of the antenna above the ground, as shown in Fig. 330.[32] The input impedance is equal

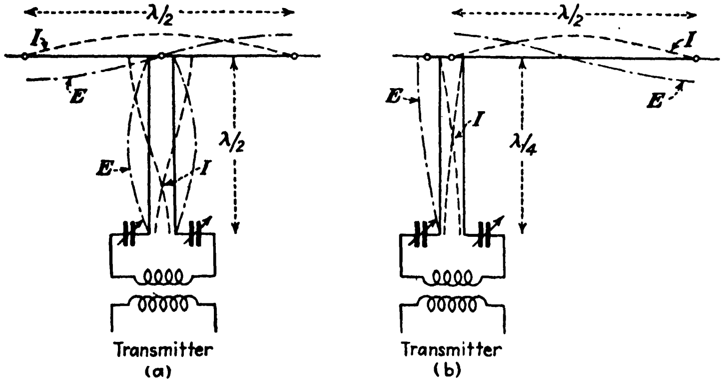


FIG. 329.—Examples of resonant transmission lines coupled to half-wave antennas.

to the terminating impedance in a resonant half-wave line because the magnitudes of E and I at the output terminals are

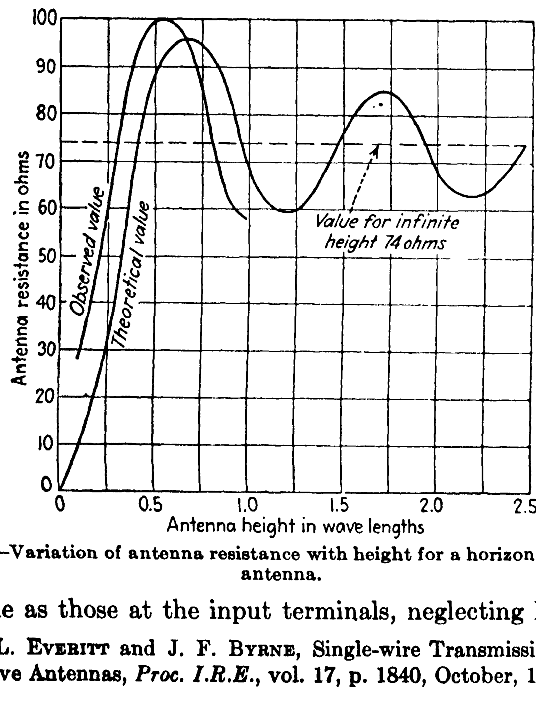


FIG. 330.—Variation of antenna resistance with height for a horizontal half-wave antenna.

the same as those at the input terminals, neglecting line losses.

[32] W. L. EVERITT and J. F. BYRNE, Single-wire Transmission Lines for Short-wave Antennas, *Proc. I.R.E.*, vol. 17, p. 1840, October, 1929.

In practice, the feeder wires are made slightly longer than the required value and are tuned to resonance by means of the two variable condensers shown. These condensers reduce the effective length of the line.

The voltage-feed system of Fig. 329*b* is commonly called a "Zeppelin" antenna. Instead of exciting the antenna at a current antinode, it is fed at one end where a current node exists and the voltage is a maximum. The required length of transmission line in this case is $\lambda/4$, or some odd multiple thereof. One wire is connected directly, while the other is capacitively coupled to the antenna.

The terms "current feed" and "voltage feed" are somewhat inaccurate in that power is supplied to the antenna in either case. In the former the current is a maximum at the connection to the antenna, while the potential there is relatively small. In the case of the voltage-feed the potential at the junction is a maximum and the current there is at a minimum.

The advantages of resonant transmission lines are that all tuning adjustments can be made at the transmitter end, and matching the impedance of the antenna to that of the line is not involved. The chief disadvantage is that the transmission line operates at practically zero power factor and produces a large resonant rise in voltage, which may introduce insulation difficulties if the power to be transmitted is large. Since standing waves are produced, radiation is avoided only by the close proximity of the two conductors. These lines are extensively used in the transmitting stations of radio amateurs. Resonant transmission lines are seldom used with receiving sets.

192. Nonresonant Transmission Lines.—The nonresonant type of radio-frequency transmission line must be terminated in an impedance equal to the characteristic impedance of the line, which involves impedance matching. When so terminated, there are no standing waves on the line and the current and voltage are practically in phase with each other at any point along the line.

The characteristic impedance is practically a pure resistance at radio frequencies equal to $\sqrt{L/C}$ for the line. The inductance of the line per unit of length is proportional to the logarithm of the ratio of the conductor spacing to the conductor diameter, while the capacitance is proportional to the reciprocal of the logarithm

of a similar ratio. The expression for the characteristic impedance of a two-wire line with sufficient accuracy for most purposes is given by

$$Z_L = 276 \log_{10} \frac{2D}{d} \text{ ohms} \tag{70}$$

where D is the conductor spacing measured between wire centers and d is the diameter of the wire, both expressed in the same units.

The expression for a concentric tube line is given by

$$Z_L = 138 \log_{10} \frac{b}{a} \text{ ohms} \tag{71}$$

where b is the inner diameter of the outer conductor and a is the outside diameter of the inner conductor. The characteristic impedance in this case is much less than with open-wire lines of the ordinary amount of spacing, which usually have a value of Z_L somewhere in the vicinity of 600 ohms.

Figure 331a shows the manner of coupling a nonresonant transmission line to a grounded antenna, such as that of a

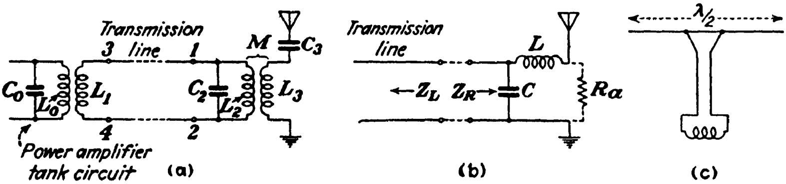


FIG. 331.—Methods of coupling a non-resonant transmission line to an antenna.

broadcasting station. The constants of the antenna circuit are reflected into L_2, so that its apparent impedance becomes, from (16), Chap. IV,

$$Z_2' = Z_2 + \frac{\omega^2 M^2}{Z_a} \tag{72}$$

where Z_2 is the vector expression for the impedance of coil L_2 with the antenna circuit open or removed, and Z_a is the vector impedance of the entire antenna circuit (including L_3) with L_2 removed. If the antenna circuit is resonated so that Z_a is a pure resistance at the radian frequency ω, only resistance will be reflected into L_2 and its inductance will remain unchanged. The amount of this reflected resistance may be varied by varying the mutual inductance M. The condenser C_2 and the impedance Z_2'

constitute a parallel-resonant circuit whose impedance as viewed from terminals 1–2 must be made equal to the characteristic impedance of the transmission line. This can be accomplished by varying L_2 and M. A line so terminated will also have an impedance of Z_L as viewed from the terminals 3–4.

Another scheme is illustrated in Fig. 331*b*. The constants of the antenna at resonance are represented by R_a, which, in series with coil L, are shunted across the condenser C. The impedance of this combination at resonance will be a pure resistance equal to

$$Z_R = \frac{L}{CR_a} \tag{73}$$

the magnitude of which will be governed by the ratio of L to C, enabling Z_R to be made equal to Z_L of the line. The parallel-resonant condition is a function of the product of L and C, so that both impedance matching and resonant operation can be secured, provided R_a is less than Z_L. If R_a is greater than Z_L, the two conditions cannot be satisfied simultaneously. If a concentric line is used, the outer sheath can be made the lower conductor in the diagram which will be at ground potential and need not be insulated from ground. The use of a tuned circuit or transformer as a coupling means is inconvenient when an ungrounded antenna is used, such as a Hertz doublet, in that the tuned element must be located in the middle of the antenna, and is therefore not readily accessible for tuning purposes. An untuned transformer of the proper ratio could be used, but the antenna would have to be operated at a frequency sufficiently below resonance so that the antenna impedance would be $R_a - jX_a$, the capacitive reactance X_a being utilized to annul the leakage reactance of the transformer. These drawbacks may be avoided at short waves by connecting the two ends of the line symmetrically to the antenna at suitable distances from the center, as Fig. 331*c*. The flared portion of the transmission line can be made to serve as coupling transformer which will be discussed later in Sec. 195.

Another form of nonresonant transmission line makes use of a single conductor.[32] No impedance-matching transformer is required in this case as the line is attached to the antenna at a point where the antenna resistance is just the value required for

[32] *Loc. cit.*

proper termination of the line. The resistance of a resonant half-wave or Hertz antenna at the center is in the vicinity of 74 ohms, whereas at either of the two ends the resistance is extremely high. This resistance is equal to the ratio of the voltage to the current at the point in question. Consequently, if the point where the line is attached to the antenna is moved outward from the center, a point P will be found displaced a distance x from the center in Fig. 332*b* where the antenna resistance is just equal to the characteristic impedance of the connecting line. Under this condition there will be no reflections from the junction point and standing waves will be absent. This is shown by the uniform distribution of current along the feeder in (*b*). When the distance x is greater or less than the proper amount, as in (*c*) and (*a*), standing waves caused by

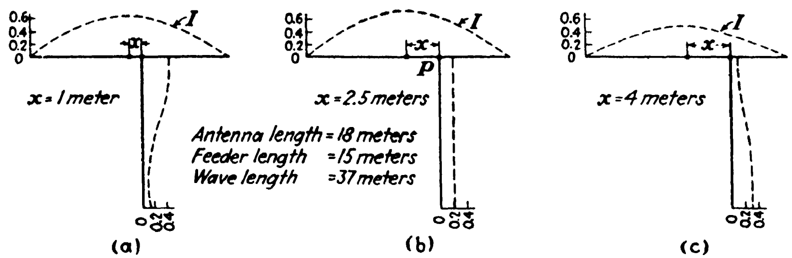


FIG. 332.—Current distribution on half-wave antenna and single-wire feeder for various feeder positions.

reflections begin to appear and appreciable radiation takes place. This will be much greater than in a two-wire line since in the latter case the radiation from the two wires tends to cancel, owing to their proximity. When properly terminated, the radiation from the usual length of feeder is small compared to that from the antenna. However, a feeder a number of waves in length may produce appreciable radiation, even when properly terminated; this being the principle of the long tilted-wire antennas discussed in Sec. 187. For transmitting purposes with a simple nondirectional antenna the radiation from the transmission line is not apt to be of considerable moment.

The chief advantage of a nonresonant transmission line over the resonant type for a given amount of transmitted power is that the losses are smaller, owing to the lower voltages and currents handled in the nonresonant case. The possibilities of radiation, or of stray pick-up in the case of reception, are much

less than with resonant lines, with the exception of the single-wire line, where faulty adjustment will make this case behave as an ordinary T antenna. For this reason a single-wire line would be unsuited for use with a highly directional receiving array. The disadvantages are that adjustments for impedance matching must be made at the junction of the line with the antenna.

193. Transmission-line Theory.—The behavior of transmission lines several wave lengths long is different in many

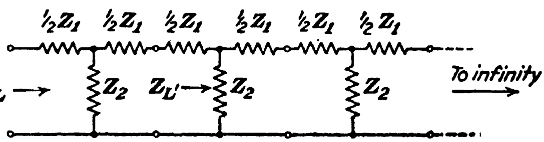


FIG. 333.—Infinite line composed of symmetrical T networks.

respects from the lines used for the commercial transmission of power at 60 cycles. The latter are electrically very short, being only a small fraction of a wave in length. The impedance viewed at the sending end is very nearly equal to the impedance of the load at the receiving end, and a short circuit at the latter end results in an excessive current in the line and generator. However, with a line which is long in the electrical sense, a short circuit at the far end may produce a very decided increase in the sending-end impedance, instead of a decrease. Some of the properties of these lines will now be investigated.

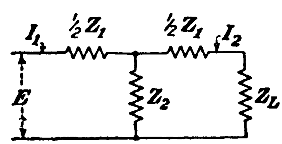


FIG. 334.—T network terminated in its characteristic impedance Z_L.

A uniform line of infinite length may be considered as being composed of a series of T networks as in Fig. 333. The *characteristic impedance* Z_L of such a structure *is defined as the vector ratio of the applied voltage to the resultant steady-state current.* Since Z_L is finite in value, the structure can absorb only a finite amount of power with a finite voltage impressed. As the total number of sections is infinite, the last or nth section can absorb only an infinitesimal amount of power, and its presence or absence can affect the value of Z_L by only an infinitesimal amount. With the nth section removed we would then have the same total number of sections in the structure as we originally had from the second section onward. Hence, the value of the impedance Z_L looking into the first section must be equal (*i.e.*, closer than any

assignable quantity) to the impedance $Z_{L'}$ with which it is terminated. Therefore $Z_L = Z_{L'}$, and the first section can be represented as shown in Fig. 334. The characteristic impedance will evidently be

$$Z_L = \frac{Z_1}{2} + \frac{Z_2\left(\frac{Z_1}{2} + Z_L\right)}{\frac{Z_1}{2} + Z_2 + Z_L}$$

$$= \sqrt{Z_1 Z_2 + \frac{Z_1^2}{4}} \tag{74}$$

The ratio of the current entering to the current leaving the section is

$$\frac{I_1}{I_2} = \frac{\frac{Z_1}{2} + Z_2 + Z_L}{Z_2} = \frac{\frac{Z_1}{2} + Z_2 + \sqrt{Z_1 Z_2 + \frac{Z_1^2}{4}}}{Z_2} \tag{75}$$

This current ratio will be the same for the other sections as well as the first, since the others are likewise terminated in an impedance Z_L.

The *propagation constant* p of the line per section (or per unit length) *is defined as the natural logarithm of the vector ratio of the steady-state currents entering and leaving the structure.*

By definition

$$p = \log_\epsilon \frac{I_1}{I_2} \tag{76}$$

or

$$\frac{I_1}{I_2} = \epsilon^p \tag{77}$$

Since I_1/I_2 will be a complex number, p must also be complex, and we may write

$$p = \alpha + j\beta \tag{78}$$

where α is the *attenuation constant* and β is the *retardation angle* or *phase constant;* both per unit length of line. Substituting (78) in (77),

$$\frac{I_1}{I_2} = \epsilon^{\alpha + j\beta} = \epsilon^\alpha \times \epsilon^{j\beta} = \epsilon^\alpha(\cos\beta + j\sin\beta) \tag{79}$$

Thus, if I_1 is the current entering the line in Fig. 335, the current I_2 at a unit distance farther along the line would be attenuated in magnitude by an amount ϵ^{α} and retarded in time phase by the angle β, the latter being due to the finite amount of time required for an electrical impulse to be propagated this distance. This angle should not be confused with the phase angle of lead or lag between the current and voltage vectors at a particular point. The locus of the current vectors will be a logarithmic spiral as shown. In the absence of attenuation the locus would be a circle. These vector relations apply to a uniform line of infinite length, but it is apparent that they will

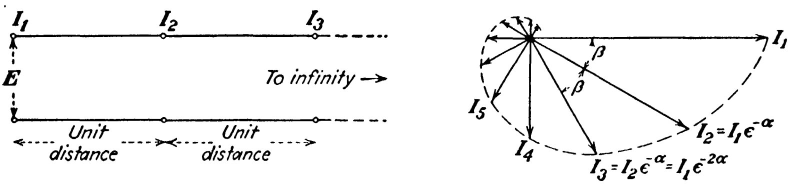


FIG. 335.—Vector relations of currents at various points along an infinite line.

also hold for the case of a finite length of line which is terminated in an impedance which is equal to the characteristic impedance of the line.

Substituting (75) in (76),

$$p = \log_{\epsilon} \frac{\dfrac{Z_1}{2} + Z_2 + \sqrt{Z_1 Z_2 + \dfrac{Z_1^2}{4}}}{Z_2} \tag{80}$$

and from the definitions of the hyperbolic functions, *viz.*

$$\sinh x = \frac{\epsilon^x - \epsilon^{-x}}{2}, \qquad \cosh x = \frac{\epsilon^x + \epsilon^{-x}}{2}, \qquad \tanh x = \frac{\sinh x}{\cosh x} \tag{81}$$

it can be shown that[33]

$$p = \cosh^{-1}\left(1 + \frac{Z_1}{2Z_2}\right) = \sinh^{-1}\frac{Z_L}{Z_2} = \tanh^{-1}\frac{Z_L}{\dfrac{Z_1}{2} + Z_2} = 2\sinh^{-1}\frac{1}{2}\sqrt{\frac{Z_1}{Z_2}} \tag{82}$$

[33] K. S. JOHNSON, "Transmission Circuits for Telephonic Communication," p. 123.

When a finite length of line is terminated in an impedance equal to Z_L, the circuit will exhibit all of the properties of a line of infinite length. There will be no electrical discontinuity at the termination and no reflections can occur at this point. It is these reflections which are responsible for the production of standing waves on the line.

194. Characteristics of Lines Having Uniformly Distributed Constants.—If the individual T sections of Fig. 333 each repre-

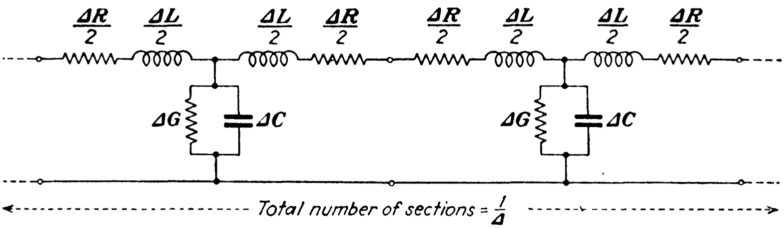


FIG. 336.—Representation of a line having uniformly distributed constants.

sent the constants of a section of line of length Δl, as in Fig. 336, the characteristic impedance is, from (74),

$$Z_L = \sqrt{Z_1 Z_2 + \frac{Z_1^2}{4}} = \sqrt{Z_1 Z_2\left(1 + \frac{Z_1}{4Z_2}\right)}$$
$$= \sqrt{\frac{\Delta(R + j\omega L)}{\Delta(G + j\omega C)}\left[1 + \frac{1}{4}\Delta^2(R + j\omega L)(G + j\omega C)\right]}$$

As Δ approaches zero, the above expression becomes

$$Z_L = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \tag{83}$$

which is the expression for the characteristic impedance of a line with uniformly distributed constants of R ohms, L henrys, G mhos, and C farads per unit of length. This unit may be any length desired as Z_L is not dependent upon the length of the line. The leakage conductance G is due chiefly to dielectric loss rather than poor insulation resistance.

From (82) the propagation constant is

$$p = \frac{1}{\Delta} 2 \sinh^{-1} \frac{1}{2}\sqrt{\Delta(R + j\omega L)\Delta(G + j\omega C)}$$

or

$$\sinh \frac{\Delta p}{2} = \frac{\Delta}{2}\sqrt{(R + j\omega L)(G + j\omega C)}$$

For small values of $\Delta p/2$ the value of sinh $\Delta p/2$ approaches $\Delta p/2$ as a limit and the propagation constant per unit of length becomes, as Δ approaches zero,

$$p = \sqrt{(R + j\omega L)(G + j\omega C)} \tag{84}$$

If the line is l units long the total propagation constant becomes lp.

Any passive structure, such as a uniform line, may be replaced at any one frequency by its equivalent T network.* In Fig. 337, let Z_1 and Z_2 be the constants of the equivalent T network of the line, which is terminated in an impedance Z_r.

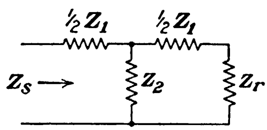


FIG. 337.—T network terminated in an impedance Z_r.

The impedance Z_s looking into the sending end of the structure will be

$$Z_s = \frac{Z_1}{2} + \frac{Z_2\left(\frac{Z_1}{2} + Z_r\right)}{\frac{Z_1}{2} + Z_2 + Z_r}$$

$$= \frac{Z_r\left(\frac{Z_1}{2} + Z_2\right) + Z_1Z_2 + \frac{Z_1^2}{4}}{\frac{Z_1}{2} + Z_2 + Z_r} \tag{85}$$

Since $Z_L = \sqrt{Z_1Z_2 + \frac{Z_1^2}{4}}$, (85) becomes

$$Z_s = \frac{Z_r\left(\frac{Z_1}{2} + Z_2\right) + Z_L^2}{\frac{Z_1}{2} + Z_2 + Z_r} = Z_L\frac{Z_r + \frac{Z_L^2}{\frac{Z_1}{2} + Z_2}}{Z_L + \frac{Z_LZ_r}{\frac{Z_1}{2} + Z_2}} \tag{86}$$

* The constants of the equivalent T network in terms of the line constants are $\frac{Z_1}{2} = Z_L \tanh \frac{lp}{2}$, and $Z_2 = \frac{Z_L}{\sinh lp}$.

From (82),

$$\tanh lp = \frac{Z_L}{\frac{Z_1}{2} + Z_2}$$

and substituting this relation in (86), we get

$$Z_s = Z_L \frac{Z_r + Z_L \tanh lp}{Z_L + Z_r \tanh lp} \tag{87}$$

which is the sending-end impedance of line having a characteristic impedance Z_L and a total propagation constant lp, when terminated in an impedance Z_r. If the propagation constant of the line is zero, tanh lp is zero and the impedance looking into the structure is equal to the terminating impedance. This is approximately the case in 60-cycle lines used for power transmission.

195. Quarter-wave-length Line as an Impedance-matching Device.—As mentioned previously, in order to connect an antenna to a nonresonant transmission line the impedance offered by the antenna should be equal to the characteristic impedance of the line. The latter, when R and G can be neglected, is given by (70) in the case of open-wire lines, and is dependent upon the ratio of conductor spacing to conductor diameter. Since this ratio is governed by practical considerations which prevent sufficiently low values of Z_L to be obtained, impedance-matching devices must generally be used to step down the value of Z_L so as to equal the much lower value of antenna resistance. At short waves it is often more economical to employ a short section of the line itself as an impedance-matching element instead of a tuned transformer.

A transmission line of length l having a characteristic impedance Z_L and propagation constant p per unit length will have an impedance at the sending end given by (87). At radio frequencies, usually $R << \omega L$ and $G << \omega C$, and p becomes, from (84)

$$p = \sqrt{-\omega^2 LC} = j2\pi f\sqrt{LC} = j\frac{2\pi}{\lambda} \tag{88}$$

where λ is in the same units as l. In other words, the attenuation constant α is negligible and $p = 0 + j\beta$. Substituting (88) in (87),

$$Z_s = Z_L \frac{Z_r + Z_L \tanh j\frac{2\pi l}{\lambda}}{Z_L + Z_r \tanh j\frac{2\pi l}{\lambda}} \tag{89}$$

If l is a half wave length, substituting $l = \lambda/2$ in (89) gives us

$$Z_s = Z_L\frac{Z_r + Z_L \tanh j\pi}{Z_L + Z_r \tanh j\pi} = Z_L\frac{Z_r + Z_Lj \tan \pi}{Z_L + Z_rj \tan \pi}$$

and since $\tan \pi = 0$,

$$Z_s = Z_r \tag{90}$$

Evidently a half-wave line is equivalent to a one-to-one transformer.

If the length of the line is $l = \lambda/4$, (89) becomes

$$Z_s = Z_L\frac{\dfrac{Z_r}{j \tan \frac{\pi}{2}} + Z_L}{\dfrac{Z_L}{j \tan \frac{\pi}{2}} + Z_r} = \frac{Z_L^2}{Z_r} \tag{91}$$

In this case a quarter-wave-length line may be used to step up or step down the value of load impedance Z_r as viewed at the

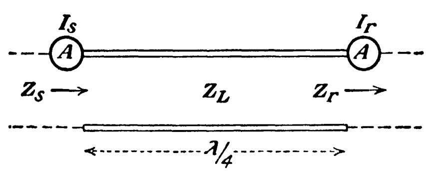


FIG. 338.—Use of a quarter-wave-length line as an impedance-matching section.

sending end, by an amount depending upon Z_L^2. This transformer action may be seen more clearly from a consideration of Fig. 338. The power entering the line will be equal to the power absorbed by Z_r, since R and G of the line are assumed to be negligible. Consequently,

$$I_s^2Z_s = I_s^2\frac{Z_L^2}{Z_r} = I_r^2Z_r \tag{92}$$

where all impedances are assumed to be pure resistances. From (92),

$$Z_r = Z_L \frac{I_s}{I_r} \tag{93}$$

$$Z_s = Z_L \frac{I_r}{I_s} \tag{94}$$

$$Z_L^2 = Z_r Z_s \tag{95}$$

where I_s/I_r is the equivalent ratio of transformation. From the above equations it follows that any two resistances which do not differ too greatly may be matched by inserting a quarter-wave line between them having the proper value of characteristic impedance.

For example, suppose a transmission line having a value of characteristic impedance equal to 600 ohms, pure resistance, is to supply energy to a load, such as an antenna, which has a resistance of 200 ohms. In order to match impedances, a quarter-wave line may be inserted between the transmission line and antenna, as in Fig. 338. The transmission line must be terminated in a 600-ohm resistance, so that $Z_s = 600$. The load offered by the antenna is $Z_r = 200$. The required value of Z_L for the quarter-wave impedance-matching section will therefore be

$$Z_L = \sqrt{600 \times 200} = 346.4 \text{ ohms}$$

If the diameter of the conductors is known, the required amount of spacing between them can be obtained from (70). Obviously, a concentric line can be employed in the same manner. Where the value of spacing required by the impedance-matching section becomes too small, the diameter of the conductors may be increased.

Another useful property of quarter-wave lines is that Z_s approaches infinity as Z_r approaches zero, if R and G of the line are negligible. A quarter-wave line short-circuited at the far end is therefore equivalent to a parallel-resonant circuit. Use is made of this interesting property at short waves in providing by-pass circuits for the low-frequency current used in melting sleet from the antenna. They are also used for draining static charges to ground, since a very high impedance is offered to the radio-frequency current and at the same time a negligible impedance is offered to direct current.

196. Transmission-line Measurements.—Equations (93) and (94) indicate that if the currents at any two points one-quarter wave length apart in a nonresonant transmission line are equal in magnitude, the line may be considered to be terminated in a load equal to its characteristic impedance. Use is made of this in practice by adjusting the load—or the impedance-matching device connected to the load—until the line currents measured at points separated by a quarter wave are alike. As it is usually impractical to insert an ammeter in the line at the various points, the current distribution along the line is ordinarily determined by an indicating device, such as a small wave meter. The deflections of the thermogalvanometer in the wave meter will be proportional to the current in the line, provided the position of the wave meter relative to line conductors is the same at each point of measurement. Suitable indicating devices are described in the paper by Sterba and Feldman.[31]

Measurements of the attenuation per unit length of the line may be made in the same manner. Expressed in decibels the loss will be

$$\text{db} = 20 \log_{10} \frac{I_s}{I_r} \tag{96}$$

where I_s is the current entering the length of line in question and I_r is the current leaving. The loss usually amounts to less than a decibel per 1000 ft. of transmission line. A 600-ohm line of this length composed of No. 6 A.W.G. (d = 162 mils) will have an attenuation of about 1 db at 40 megacycles. At lower frequencies or with larger conductors, the attenuation will be correspondingly lower. With a concentric line the attenuation is a minimum for a fixed diameter of outer conductor when the ratio of the outer to the inner diameters is

$$\frac{b}{a} = 3.6 \tag{97}$$

From (71), this optimum ratio results in a characteristic impedance of 77 ohms, which is considerably lower than with open-wire lines.

197. Transmission Lines for Reception.—For point-to-point communication using a fixed frequency the transmission line

[31] *Loc. cit.* p. 474.

connecting the antenna to the receiving unit differs but little from the transmitting types, except in the size and spacing of the conductors and the additional precautions which must be taken to avoid extraneous pick-up. The conventional T and inverted-L antennas commonly used for broadcast reception are apt to pick up a large amount of noise on short waves when an all-wave receiver is used. A great deal of this is due to pick-up by the lead-in wire by virtue of its greater exposure to lighting circuits and other electrical wiring in the building. These circuits are capable of transmitting and radiating electrical disturbances that are produced by various electrical devices located some distance away from the receiver. If the antenna lead-in

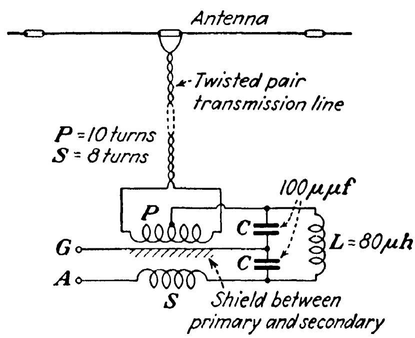


FIG. 339.—Antenna coupling device used with all-wave receiving sets.

is replaced by a suitable transmission line which is properly transposed, equal voltages will be induced in the two wires and will cancel. The antenna, being farther removed from the disturbing circuits, is affected to a lesser extent than the lead-in. Its horizontal position also reduces the pick-up of vertically polarized radiation. A fair amount of "man-made" static is vertically polarized.

A twisted pair of conductors, similar to lamp cord, is the usual form of transmission line used for short-wave-broadcast reception. The chief difficulty is in the matter of proper termination, as an impedance match cannot be maintained over the entire frequency range of an all-wave receiver by means of a single transformer of fixed ratio. This requirement is not so important with reception, however, as the average receiving set has considerably more sensitivity than can be usefully employed. The coupling device is usually designed to give proper termination

when used with a doublet antenna of fixed length for the short-wave bands where desirable programs may be heard. The performance at other wave lengths is correspondingly poorer.

A horizontal Hertz doublet is not so efficient on the broadcast frequencies as a vertical antenna, since the ground wave produced by these stations is vertically polarized. A coupling device which causes a gradual transition in antenna operation from a T antenna at broadcast wave lengths to a doublet antenna at short waves is shown in Fig. 339. At short waves the receiver is transformer-coupled to the transmission line, while at the longer waves two transmission-line conductors are tied together by the transformer primary, the mid-point of which is connected to the antenna terminal of the set through the coil L and the transformer secondary. The reactance of these two items in series is small at broadcast frequencies, but becomes very much larger at high frequencies. A grounded shield is placed between the two transformer windings to prevent electrostatic coupling.

198. Propagation of Electromagnetic Waves.—The successful reception of signals transmitted across the Atlantic by Marconi in 1901, needed some explanation in that these waves, if similar to light, should be unable to follow the curvature of the earth. Heaviside in England and Kennelly in the United States offered the suggestion that an ionized layer in the upper atmosphere could serve as a reflecting surface which would confine the radiation to the earth. This conducting layer has been given the name "Kennelly-Heaviside layer." Subsequent investigation, particularly in recent years, has led to considerable modification in the original hypothesis. The existence of not one, but of several ionized layers has been demonstrated by experiments. Instead of the wave being reflected from the conducting layer, as light from the surface of a mirror, it enters the medium and is bent back to earth again by refraction. However, it is convenient for purposes of calculation to regard the process as one of simple reflection whereby the wave travels with the velocity of light to the hypothetical reflecting plane, where it is then reflected and returned at the same velocity. The height of this plane is called the virtual height of the layer, which is seen from Fig. 340 to be somewhat greater than the actual height.

When the wave enters the ionized medium, the signal or group velocity is retarded below the velocity of light. This group

velocity is the rate at which the energy of the wave is traveling. The phase velocity of the wave is the rate at which the phase changes along the path of the wave and can be much greater than the velocity of light, since it is equal to the velocity of light divided by the refractive index of the medium. This speeding up of the phase velocity causes the wave path to be bent back to the earth. The amount of bending experienced will depend upon the degree of ionization and its gradient, as well as upon the frequency of the incident wave. With a given density of ionization the amount of bending will diminish with an increase of frequency, so that at some critical value of the latter the wave will penetrate the lowest ionized layer and be reflected from some more strongly ionized layer above the first. Wave lengths below 8 or 10 meters will usually not be bent sufficiently to return to earth, and hence are not useful for long-distance communication.

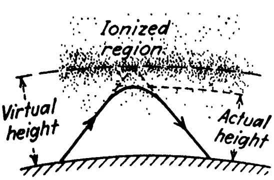


FIG. 340.—Refraction of a radio wave by an ionized medium.

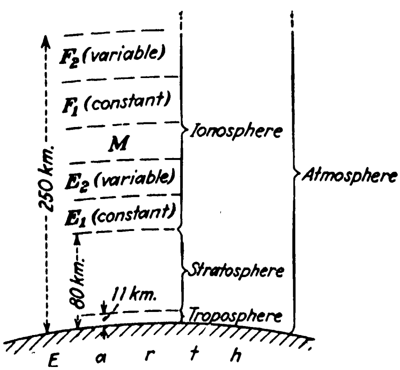


FIG. 341.—Location of various ionized layers in the upper atmosphere.

199. The Ionosphere.[34]—The atmosphere surrounding the earth may be divided into three or more layers, as shown in Fig. 341. The lowest of these, extending upward to about 10

[34] For an historical summary which outlines the principal published reports of ionosphere studies, together with their own observations, the reader is referred to a paper by S. S. Kirby, L. V. Berkner, and D. M. Stuart, Studies of the Ionosphere and Their Application to Radio Transmission, *Proc. I.R.E.*, vol. 22, p. 481, April, 1934.

or 12 kilometers above the earth is called the troposphere. Clouds form in this region and the temperature decreases with the height. Above this lies the stratosphere wherein the temperature no longer varies with the altitude. At a height above 80 kilometers appreciable ionization begins, suggesting the name ionosphere for this region. The ionosphere may be regarded as being composed of two general regions, the lower or E region, and a more intensely ionized F region having a height of from 200 to 300 kilometers. The latter is usually composed of two distinct layers known as F_1 and F_2, the virtual heights of which may

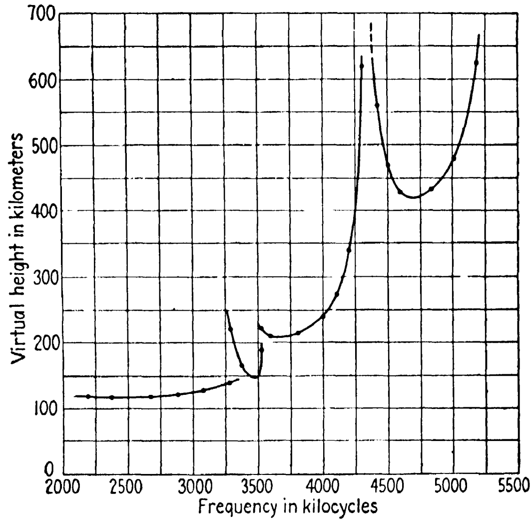


FIG. 342.—Variation of virtual height with signal frequency.

change abruptly. Occasionally these two layers merge into a single layer. Two distinct layers are also observed at times in the E region.

The measurements of the virtual heights of these layers are made by measuring the time required for a signal to travel to the reflecting layer and back to earth. Short pulses are sent out by the transmitter at some convenient rate such as 60 per second. The direct and reflected impulses are picked up by a receiving set and impressed upon a suitable oscillograph, such as the cathode-ray type. By measuring the time interval between the direct impulse and its "echo" the virtual height may be determined. A typical record reported by Goodall[35] is shown in Fig. 342. As

[35] W. M. GOODALL, The Ionosphere, *Bell. Lab. Rec.*, vol. 13, p. 194, 1935.

the frequency is increased, the virtual height remains practically constant until the critical value is reached where penetration of the reflecting layer occurs. The critical frequency for the *E* layer varies from 2500 to 3500 kc, depending upon the season of the year and time of day. Frequencies below the critical value are all reflected by this layer, but with increased attenuation as the critical frequency is approached. As penetration begins, there is usually an abrupt increase in the virtual height and reflections from both layers may be observed. The virtual

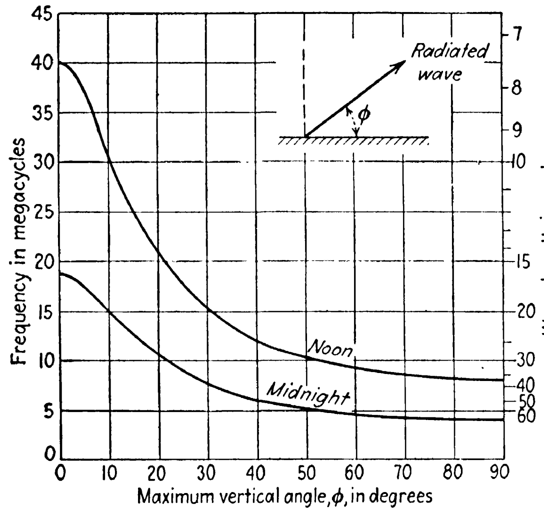


FIG. 343.—Curves showing average values of maximum vertical angle for which a wave of given frequency can be reflected from the ionosphere.

height beyond the critical frequency is that of the next higher layer. The large increase in virtual height at the critical frequency is due to the decrease in the velocity of propagation through the penetrated layer.

The frequency at which penetration occurs will also depend upon the angle of incidence which the ray or beam of radiation makes with the reflecting layer, the data of Fig. 342 being for waves of normal incidence. The greater the angle of incidence, the shorter the wave length required to penetrate completely all of the reflecting layers. This is shown in Fig. 343 for average noon and midnight conditions in terms of the angle ϕ which the radiated beam makes with the horizontal. Frequencies lying above the curves will not be bent sufficiently to return to earth,

and hence cannot be used for transmission purposes for distances beyond the range of the ground wave.

The ionizing agent responsible for these layers is probably ultraviolet light from the sun, perhaps augmented by some form of corpuscular bombardment. In the lower layers where the gas pressure is higher, ionization and recombination closely follow the illumination by the sun. The much lower pressure in the higher layers results in much slower rate of recombination and the gas remains ionized long after sunlight is removed.

200. Skip-distance Effect.—The amount of attenuation experienced by the ground wave increases with the frequency, while the amount of bending produced by the ionosphere becomes less. As a result, for frequencies above 6000 kc the ground wave is

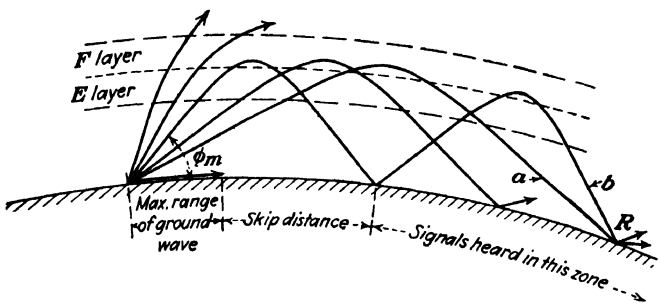


FIG. 344.—Propagation of short waves, illustrating skip-distance effect. A ray having a vertical angle greater than ϕm will not return to earth.

attenuated to a value below useful audibility at a moderate distance from the transmitter before the sky wave is reflected to earth. This produces a skip-distance zone where no signal at all is heard, as illustrated in Fig. 344. The high-angle radiation, which, at broadcast frequencies (at night), is reflected to earth a short distance from the transmitter, at high frequencies now penetrates the reflecting layer and travels into space. The width of the skip-distance zone increases with the frequency, becoming greater than 1000 miles in daylight at 30 megacycles. The skip distance is greater at night than under daylight conditions and is greater in winter than in summer. Sporadic signals are occasionally heard within this zone, owing to abnormal conditions and scattering. The latter is similar to the diffused light from a rough surface when illuminated by incident light. With high power at certain frequencies signals may be heard in this zone

which have traveled completely around the world, or else have traveled in the opposite direction and reached the receiver by an indirect great-circle route. This latter mode of reception is possible only when a nondirectional antenna is used for transmission.

201. Echoes.—Around-the-world signals of the types just mentioned, when the receiver is beyond the skip distance, give rise to the reception of multiple signals, referred to as "echoes." The first signal impulse is received by direct transmission from the source, to be followed approximately $\frac{1}{7}$ sec. later by a weaker repetition of the same impulse which in this time has made a complete trip around the earth. On occasions, two or more repetitions of the original signal have been observed which represent two or more complete trips around the world in addition to the distance separating the transmitting and receiving stations.[36] Multiple signals of this type are observed only at times when the great circle between transmitter and receiver lies along the twilight zone, and for wave lengths from 18 to 20 meters. Under these conditions the bending of the wave by the ionosphere is accompanied by very low attenuation. Multiple signals of shorter delay will be produced by signals traveling from the transmitter by the indirect route away from the receiver when nondirectional antennas are used.

A second type of multiple signal of very brief delay is caused by the received signals traveling by paths having different lengths, as in Fig. 344. The beam *a* arrived at point *R* by a shorter path, involving only a single reflection, than did the one following a path such as *b*, which involved several reflections. Multiple signals of this type are not particularly serious when moderate speeds of transmission are used as the time delay is not sufficient to cause confusion, although it may at high speeds. With radio telephony these multiple signals produce a characteristic hollow sound as though the speaker were talking into a microphone located in an empty room. The reverberations of the sound waves in the room are exactly analogous to the similar case of the radio waves undergoing successive reflections between earth and ionosphere. At great distances the "reverberation time" will be small, as the waves are attenuated at each reflection during

[36] E. QUACK and H. MÖGEL, Double and Multiple Signals with Short Waves, *Proc. I.R.E.*, vol. 17, p. 791, May, 1929.

the passage through the E layer. Consequently, only the beam which arrived by a path involving the minimum number of reflections is apt to be heard.

Multiple signals are a serious nuisance in the facsimile transmission[37] of pictures and printed matter, as a double image of the transmitted material is produced at the receiver, the two (or more) images being displaced from each other by an amount depending upon the speed of transmission and the time interval between the multiple signals.[38] These multiple signals limit the speed of transmission to values slow enough to avoid blurred images. Similar effects would be produced in television images if these same wave lengths and distances of transmission were used. For these and other reasons wave lengths in the vicinity of 5 to 7 meters are more suitable for television purposes.

A third type of echo of very great delay has been occasionally observed, ranging from several seconds to over 4 min., indicating the wave has either suffered a great reduction in group velocity, or else has traveled an enormous distance. These are a matter of scientific rather than engineering interest and the various theories advanced in explanation will not be considered here. A discussion of them will be found in the reference below.[39]

202. Fading.—One form of fading that has been previously mentioned is due to the reflection of the sky wave to earth where it combines with the ground wave at that point. The resultant field strength will be their vector sum and will depend upon the phase displacement between the two components. If the lengths of the two paths are both an even number of half wave lengths the two waves will reinforce each other, but if the length of one path measured in half-waves is odd, while the other is even, they will be 180 degrees out of phase and will tend to annul each other. As the character of the ionosphere is continually changing the length of the path traversed by the sky wave will likewise

[37] For a description of such systems see R. H. Ranger, Transmission and Reception of Photoradiograms, *Proc. I.R.E.*, vol. 14, p. 161, April, 1926, and Mechanical Developments of Facsimile Equipment, *ibid.*, vol. 17, p. 1564, September, 1929. Also V. Zworykin, Facsimile Picture Transmission, *ibid.*, vol. 17, p. 536, March, 1929.

[38] T. L. Eckersley, Multiple Signals in Short-wave Transmission, *Proc. I.R.E.*, vol. 18, p. 106, January, 1930.

[39] P. O. Pedersen, Wireless Echoes of Long Delay, *Proc. I.R.E.*, vol. 17, p. 1750, October, 1929.

vary, causing the field strength at the receiver to rise and fall. This form of fading is quite common at broadcast frequencies at night, and may be very serious at distances of only 50 miles from the transmitter. Fading of this type is absent during the daytime, as the sky wave is almost completely attenuated by the ionosphere in this range of frequencies. Fading of this type cannot be produced by waves short enough to produce a skip-distance effect.

Fading beyond the range of the ground wave is caused by interference due to phase differences between two or more sky waves, such as *a* and *b* in Fig. 344. Even with only a single beam, changes in the amount of attenuation experienced upon reflection would produce fluctuations in the received field strength. Likewise, changes in the plane of polarization would cause similar effects. Short-wave reception is characterized by more rapid and continual changes in signal strength than reception at longer waves.

Another form, known as selective fading, produces very serious distortion, in the case of modulated signals, in addition to changes in volume. This is caused by carrier and side-band frequencies traveling by paths differing in length, which results in phase shifts between the components of the signal. In addition, frequencies differing by only a few hundred cycles have been found to fade in and out independently of each other, so that it is possible at times for the carrier frequency to disappear and the two side bands to survive, making the result completely unintelligible.[40] This form of fading at broadcast frequencies is present only at night. Reception in the daytime is usually limited to the range of the ground wave which does not follow multiple paths.

Very severe fading is frequently encountered at short waves when the receiving station is near the edge of the skip-distance zone. A minor change in the ionosphere may cause the waves to pass completely over the station. At times of magnetic storms, which are usually associated with sun-spot activity, short-wave signals may fade out completely for several hours. As solar activity varies through an 11-year cycle, some types of

[40] Fading at broadcast frequencies is discussed in an excellent paper by R. Bown, D. K. Martin, and R. K. Potter, Some Studies in Radio Broadcast Transmission. *Proc. I.R.E.*, vol. 14, p. 57, February, 1926.

short-wave-transmission phenomena may be expected to follow a similar cycle. For example, wave lengths which were found to be usable at night for transmission between certain fixed points in 1928 proved to be entirely too short for the same service five years later.

203. Long-wave Transmission.—The lower the frequency of transmission, the less the attenuation of the ground wave. Most of the received energy up to distances of perhaps 500 miles is by means of the ground wave at wave lengths above 10,000 meters, and diurnal and seasonal effects are absent within this distance. At greater distances the sky wave begins to predominate and maximum signal strength is obtained when the transmission path lies in complete darkness. As soon as the sunset line falls across the transmission path, there is practically always an abrupt reduction in field strength, followed shortly by a rise to normal night values. The differences between day and night field strengths become more marked as the wave length is reduced to medium values, and the dip at sunset becomes more pronounced. The differences between summer and winter reception are not very great at long waves. At the higher latitudes the time the entire path is in complete darkness is much shorter in summer than in winter. Static is also more severe in the summer, and stronger at night than during the day. At night the disturbances are propagated farther owing to decreased absorption.

The amount of energy absorbed from the sky wave by the ionosphere becomes less as the wave length is increased, so that for communication at great distances as long a wave length as possible should be used. Unfortunately, the radiation efficiency of the antenna is low at these wave lengths, and directive transmission is impractical because of the prohibitive size of antenna structure that would be required.

204. Critical Wave Lengths.—The gas pressure in the lower portion of the E layer results in an average collision frequency of electrons with gas molecules of about 1.4×10^6 cycles per second. Waves of this frequency will therefore experience the maximum amount of absorption, while waves of lower frequency will be reflected with correspondingly less attenuation. Higher frequencies will pass through the layer and will be attenuated to some extent, but to a lesser degree than those in the vicinity of the critical frequency. Waves ranging in length from about 150

to 500 meters will therefore be absorbed to a greater extent than the longer or shorter waves, and they may be regarded as the transition band between long- and short-wave-transmission phenomena. They are the least suited for long-distance-transmission purposes and are used chiefly for broadcasting.

As previously mentioned, satisfactory broadcast coverage is obtained only in the area where the ground wave is strong enough to override static and local interference. The return of the sky wave to earth at night produces serious fading at nominal distances where the two waves are of comparable intensities. Since the attenuation of the ground wave increases with the frequency, the lower broadcast frequencies will result in larger service areas for a given amount of radiated power. In addition to the frequency, the attenuation also depends upon the character of the country over which the ground wave passes. It is a minimum over water, or flat country of good soil conductivity, and somewhat higher over broken country. The tall steel-framed buildings of cities with their overhead wiring and large built-up areas produce a very high degree of attenuation.[41]

205. Short-wave Transmission.[42]—One of the reasons why the advantages of short-wave transmission for long distances remained undiscovered until about 1922 was that extensive investigations which had been made on long-distance transmission all showed that the attenuation increased rapidly as the frequency was raised, especially for daylight transmission. Any sporadic transmission experiments using short waves that may have been made at various times to determine their range would have encountered the inner boundary of the skip-distance zone, and existing theories would have seemingly been confirmed. Wave lengths of 200 meters and below had been assigned in the United States to amateur radio operators. As considerable congestion existed in the immediate vicinity of 200 meters, some of the more adventurous individuals began exploring somewhat shorter wave lengths with good results. A few research laboratories also began

[41] See papers by L. ESPENSCHIED, Radio Broadcast Coverage of City Areas, *Trans. A.I.E.E.*, vol. 45, p. 1278, 1926; C. M. JANSKY, JR., Some Studies of Radio Broadcast Coverage in the Mid-west, *Proc. I.R.E.*, vol. 16, p. 1356, October, 1928.

[42] For experimental data on the variation of field strength with distance see C. N. Anderson, Notes on Radio Transmission, *Proc. I.R.E.*, vol. 19, p. 1150, July, 1931.

similar experiments using amateur observers as listeners. The data obtained soon indicated that the propagation of short waves followed an entirely different law from the long waves, and commercial short-wave stations began to increase rapidly in number.

The choice of wave length for communication over a given distance will depend upon a number of factors.[43] The shortest wave length that will be reflected to the earth within the required distance should be selected. Longer wave lengths could be employed, but these would be reflected to earth much sooner and would require several reflections between earth and ionosphere to make the trip, with the attendant increase in attenuation at each reflection. When the distance is very great and the entire path is in daylight, a wave length in the vicinity of 15 meters would be satisfactory. For a lesser distance the skip distance at 15 meters might be too great, necessitating the use of a longer wave. At night these wave lengths would be too short to be reflected to earth and a wave in the vicinity of 30 to 40 meters would probably be used. At twilight, or when the transmission path is partly in daylight and partly in darkness an intermediate wave length between the day and night values must be used. Conditions will also differ somewhat from winter to summer, but by suitable compromises satisfactory 24-hr. communication throughout the year can usually be secured by the use of only three frequencies. When uninterrupted communication must be insured at all times, additional frequencies should be provided.

Wave lengths below 8 or 10 meters are not reflected and hence are usable only for special short-distance applications such as blind-landing beams for aviation, police communications, telephone extensions to near-by islands, etc. These waves are refracted with a radius of curvature greater than that of the earth, so that the sky wave travels out into space. The transmission distance for these quasioptical waves would be theoretically the visual distance between the transmitter and receiver, although the longer waves of this group can usually be received at locations somewhat below the horizon owing to diffraction

[43] An extensive discussion of these is given by M. L. Prescott, The Diurnal and Seasonal Performance of High-frequency Radio Transmission, over Various Long-distance Circuits, *Proc. I.R.E.*, vol. 18, p. 1797, November, 1930.

and scattering effects.[44] For point-to-point service the transmitting and receiving antennas are placed as high as possible. as on the roof of a tall building, or on the top of a high hill. Telephone service between the islands of the Hawaiian group by means of radio links has been in commercial operation since 1932, using wave lengths from 5 to 10 meters. An experimental circuit operating on 220 to 230 megacycles (1.3 meter) over an optical path of 56 miles was installed in the spring of 1935.[45]

Wave lengths below 2 meters, when the transmission distance permits their use, are very satisfactory for point-to-point communication as they are practically immune from fading, static, and automobile ignition interference. The directional antennas are small and relatively inexpensive, and the transmitted beam can be made to approximate that of a searchlight, insuring the complete privacy of the conversation.

Problems

1. What is the value of the field strength in volts per meter at the earth's surface at a point 1 mile from a vertical quarter-wave antenna when the current in the base of the antenna is 10 amp.? What would it be for an airplane 5000 ft. above this point?

2. If the operating frequency of the above antenna is increased so that $h/\lambda = 0.6$, the current at an antinode remaining at 10 amp., what will be field strength at the same points as in Problem 1?

3. At what finite elevation of the airplane in Problem 2 would the field strength be zero?

4. Plot a curve showing the energy distribution in the first quadrant around a vertical antenna operated so that $h/\lambda = 0.625$. On the same area for comparison purposes show the energy distribution around a quarter-wave antenna. Assume the factor $60I_m/r_0$ to be unity in both cases.

5. A loop 3 ft. square has 10 turns and has a value of Q equal to 40 at 5×10^5 cycles. When resonated to this frequency and rotated for maximum induced e.m.f., the voltage across the tuning condenser is 4 millivolts. What is field strength of the received signal in millivolts per meter?

6. Plot a polar diagram of the relative field strength in the horizontal plane around two vertical transmitting antennas excited in phase and separated by a distance of $\frac{3}{8}\lambda$.

7. Repeat Problem 6 with the current in the right-hand antenna lagging $\frac{1}{8}$ period behind the current in the left-hand antenna.

[44] C. R. BURROWS, A. DECINO, and L. E. HUNT, Ultra-short-wave Propagation Over Land, *Proc. I.R.E.*, vol. 23, p. 1507, December, 1935; also E. O. HULBERT, The Ionosphere, Skip Distances of Radio Waves, and the Propagation of Microwaves, *ibid.*, p. 1492.

[45] W. I. HARRINGTON and C. W. HANSELL, The Hawaiian Radio Telephone System, *Elec. Eng.* vol. 54, p. 822, August, 1935.

8. The antenna system of station WOR at Carteret, N. J., consists of three vertical radiators excited in phase and spaced 395 ft. apart. The operating frequency is 710 kc. Plot a polar diagram of the relative field strength in the horizontal plane. The major axis of the curve lies along a line from New York to Philadelphia, the latter being 60 miles from Carteret. If the field strength at Philadelphia due to the ground wave is 4 millivolts per meter with 50 kw supplied to the antenna, what would be the field strength due to the ground wave an equal distance from the antenna but at right angles to the major axis? Assume the attenuation along both paths to be equal.

9. A radio-frequency transmission line composed of No. 4 AWG. conductors is to have a characteristic impedance of 600 ohms. What must be the spacing between conductors? It is desired to replace this line with a concentric line of the same impedance having an inside diameter of the outer conductor of 2.5 in. What must be the diameter of the inner conductor? Would this be practical?

10. An antenna is coupled to a 600-ohm transmission line by means of the circuit of Fig. 331*a*. The antenna has a resistance of 120 ohms at the operating frequency of 10^6 cycles, the resonant adjustment being made by means of L_3 with L_2 removed. If C_2 is 0.001 μf, what must be the values of M and L_2 for proper termination of the line?

11. A concentric transmission line, which has an inner-conductor diameter of 0.5 in. and an inner diameter of outer conductor of 2 in., is to be connected to an antenna having a resistance of 40 ohms by means of the circuit of Fig. 331*b*. Find the values of L and C required for proper termination of the line.

12. Show that the termination of Fig. 331*b* can be employed only when the characteristic impedance of the line is greater than the resistance of the antenna.

13. A transmission line composed of No. 6 AWG. conductors spaced 1 ft. apart has a length of $\frac{3}{4}\lambda$. Assuming negligible losses, what is the vector expression for the impedance looking into the line when the far end is short-circuited? When the far end is open-circuited?

14. The constants per foot of a twisted-pair transmission line for reception are: $L = 0.1726$ μh, $C = 18.2$ $\mu\mu$f. The power factor due to dielectric loss is 1.62 per cent and is practically constant for all frequencies. If the conductor resistance can be neglected what is the attenuation in decibels per 1000 ft. at 3.5 megacycles? At 1.592 magacycles?

15. In Problem 14, if the circuit resistance at 3.5 megacycles is 46.8 ohms per 1000 ft. (two wires), what will the attenuation be at this frequency?

16. A transmission line of the same construction as Problem 13 is to be connected to a 180-ohm antenna array by using a quarter-wave line as an impedance-matching device. If the operating frequency is 20 megacycles, specify the length and spacing of the quarter-wave section, assuming the conductor diameter to be the same as the transmission line.

17. In Problem 16, if the spacing between conductors of the quarter-wave section must remain the same as for the transmission line, what diameter conductors must be used to match impedances?

NAME INDEX

A

Adcock, F., 443
Affel, H. A., 331
Alexanderson, E. F. W., 318, 381, 474
Anderson, C. N., 500
Armstrong, E. H., 117, 217, 374, 382
Arnold, H. D., 60, 116
Ataka, H., 374

B

Bailey, A., 467
Bailey, S. L., 449
Ballantine, S., 360, 392, 429, 432, 433, 434
Barber, I. G., 61
Barkhausen, H., 309
Barton, L. E., 195
Beck, A. C., 471
Becker, J. A., 123
Bellini, A., 441
Berkner, L. V., 492
Beverage, H. H., 465
Black, H. S., 218, 220
Blackwell, O. B., 331
Bown, R., 334, 498
Brainerd, J. G., 348
Brown, G. H., 417
Bruce, E., 468, 471
Burrows, C. R., 502
Bush, V., 129
Byrne, J. F., 476

C

Cady, W. G., 294
Carson, J. R., 334, 337, 354, 429
Carter, P. S., 428
Chaffee, E. L., 115, 142
Child, C. D., 119
Christopher, A. J., 250
Clapp, J. K., 304
Colpitts, E. H., 268, 331

D

Dean, S. W., 467
Decino, A., 502
De Forest, Lee, 116, 117, 344
Dellinger, J. H., 443
Demarest, C. S., 331
Diamond, H., 240, 443, 448, 449
Donisthorpe, H., 442
DuBridge, L. A., 121
Dunmore, F. W., 443

E

Eckersley, T. L., 418, 497
Edison, T. A., 116
Eglin, J. M., 211
Elder, F. R., 308
Elmen, G. W., 60, 63
Elster, J., 115
Espenschied, L., 500
Evans, H. P., 338
Everitt, W. L., 283, 476

F

Fay, C. E., 285
Feldman, C. B., 474, 489
Fleming, J. A., 116
Foster, R. M., 452
Franklin, C. S., 117
Fuller, L. F., 306

G

Geitel, H., 115
Gihring, H. E., 417

Gill, E. W., 310
Glessner, J. M., 370
Goodall, W. M., 493
Googin, T. M., 353
Green, C. W., 331

H

Hanna, C. R., 67
Hansell, C. W., 428, 502
Harrington, W. I., 502
Harris, W. A., 388, 406
Hartley, R. V. L., 266
Hazeltine, L. A., 53, 256
Heaviside, O., 491
Heising, R. A., 319, 334
Herold, E. H., 388
Hertz, H., 413, 436
Hoare, S. C., 377
Hollmann, H. E., 309
Horton, J. W., 300
Housekeeper, W. G., 117
Hughes, A. L., 121
Hulbert, E. O., 502
Hull, A. W., 116, 128, 304, 307
Hull, L. M., 304
Hund, A., 297
Hunt, L. E., 502

I

Iinuma, H., 305

J

Jackson, W. E., 449
Jansky, C. M., Jr., 500
Johnson, K. S., 222, 357, 483

K

Kear, F. G., 449
Keith, C. R., 333
Kellogg, E. W., 411, 465
Kennelly, A. E., 491
Kerr, G. P., 211
Kilgour, C. E., 182, 370
King, R. W., 139
Kirby, S. S., 492
Kishpaugh, A. W., 327
Kreer, J. G., 218
Kurz, K., 309
Kusunose, Q., 139

L

Lack, F. R., 296
Langmuir, I., 116, 120
Langsdorf, A. S., 419
Lawrence, R. R., 455
Levin, S. A., 428
Lindenblad, N. E., 428
Llewellyn, F. B., 301, 338, 341, 348
Loftin, E. H., 211
Loughren, A. V., 181
Lowry, L. R., 471

M

McArthur, E. D., 308
McDonald, W. A., 246
McIlwain, K., 348
McNally, J. O., 192
McPetrie, J. S., 457
Marconi, G., 413, 491
Marrison, W. A., 296
Martin, D. K., 334, 498
Mason, F., 298
Massa, F., 412
Mathes, R. C., 400
Maxwell, J. C., 413
Megaw, E. C., 309
Meissner, A., 117, 268
Mendenhall, H. E., 282
Mitchel, D., 402
Mögel, H., 496
Morgan, N. R., 358
Morrell, J. H., 310

N

Nesslage, C. F., 388
Nyquist, H., 218

O

Okabe, K., 310
Olsen, H. F., 412

P

Pedersen, P. O., 306, 497
Peterson, E., 218, 333, 338
Pierce, G. W., 82, 229, 427
Polydoroff, W. J., 56
Potter, R. K., 334, 498
Poulsen, V., 306
Pratt, H., 443, 445
Prescott, M. L., 501
Prince, D. C., 273

Q

Qüack, E., 496

R

Ranger, R. H., 497
Rice, C. W., 254, 411, 465
Richardson, O. W., 118
Roder, H., 317
Round, H. J., 117

S

Schmitt, O. H. A., 211
Schottky, W., 116
Shackelton, W. J., 61
Smith, C. G., 129
Smith-Rose, R. L., 439
Snow, H. A., 392
Solt, C. T., 442
Southworth, G. C., 453, 457
Spitzer, E. E., 308
Sterba, E. J., 459, 462, 474, 489
Stowell, E. Z., 240
Stuart, D. M., 492

T

Taylor, A. H., 211
Terman, F. E., 353, 356, 358
Thompson, B. J., 187, 310
Thomson, J. J., 116
Thuras, A. L., 410
Tosi, A., 441

U

Uehling, E. A., 93

V

Van der Bijl, H. J., 330
VanDyke, K. S., 297

W

Walmsley, T., 459
Ware, L. A., 218
Warner, J. C., 181
Wehnelt, A., 122
Wente, E. C., 410
Wheeler, H. A., 49, 246, 385
White, S. Y., 211
White, W. C., 308
Williams, N. H., 116
Wilmotte, R. M., 457
Wilson, E. D., 121
Wintringham, W. T., 467
Woods, F. S., 337
Wright, S. B., 400, 402

Y

Young, C. J., 428

Z

Zottu, P. D., 310
Zworykin, V. K., 121, 497

SUBJECT INDEX

A

A.V.C., 394–398
Absorption modulation, 318, 319
Acoustic feed-back, 213
Acoustic tests on receiving sets, 410
Acoustically compensated volume control, 402, 403
Activation of filament, oxide coated, 123
  thoriated tungsten, 124
Adcock antenna, 443
Admittance, 14
  grid input, 228–235
Aerial (*see* Antenna)
Aircraft, radio beacons for, 443–449
Alloys, magnetic, 60–64
Alternating-current power, 4, 19
Amplification, expressed in decibels, 170
  graphical determination of, 152
  measurement of, 171, 172
  regenerative, 212, 213
Amplification factor, 131
  measurement of, 141
Amplifiers, acoustic feed-back in, 213
  approximate calculations of output, 183, 190
  audio-frequency, 144–222
  classification of, 145
  comparison of, audio frequency, 169
  Class A, 145
  Class AB, 200, 201
  Class B, 145, 195–200, 287
  Class C, 145, 289
  direct-coupled, 209–211
  direct-current, 209–211
  distortion in, 146–150
  double-grid, 198, 199
Amplifiers, double-impedance-coupled, 160
  feed-back in, 194, 212–220
  filtering of, 216
  graphical determination of distortion in, 182, 187, 208
  hum in, 191
  impedance-coupled, 157–161
  with inductive load, 161
  input capacitance of, 154, 228–235
  Loftin-White, 211
  maximum undistorted output of, 180, 181
  microphonic effects in, 170, 213
  modulated Class C, 290
  neutralization of, 254, 259, 293, 294
  parallel-feed, 161
  power, Class A, 177–179, 184–190, 204–209
    conditions for maximum output, 180, 190
    distortion, determination of 182, 187, 208
    double-grid, 198, 199
    graphical determination of output, 179, 189, 208
    pentodes used as, 204–209
    (*See also* Power amplifiers)
  prevention of feed-back in, 215–217
  push-pull, 184–192, 195–201
    advantages of, 189, 190
    Class AB, 200, 201
    Class B, 195–200
    distortion in, 187
    output of, 189, 190
  radio frequency in cascade, 251, 252
    effect on selectivity and fidelity, 252

Amplifiers, radio-frequency, for reception, 237–259
for transmission, 269–294
(*See also* Power amplifiers, radio-frequency)
regeneration in, 212–220
resistance-coupled, 154–157
with resistance load, 150
reversed feed-back, 218–220
self-bias, 192–195
space-charge grid, 204
tetrode, 201–204
transformer-coupled, 163–169
characteristics, 167
equivalent circuit, 165
tuned, 239–252
effect of mutual inductance, 243–246
inductive and capacitive coupling in, 246–249
neutralization of, 253–258
optimum coupling in, 242
parallel resonant, 239
selectivity of, 245
transformer-coupled, 239–243
using tuned coupled circuits, 249–251
untuned, 237–239
Amplitude distortion, 146
Amplitude modulation, 315–317
methods of obtaining, 318
Antennas, 416–474
Adcock, 443
array, 429
gain of, 462
polar diagrams of, 454–457
practical aspects of, 461–463
principles of, 449–454
types of, 458–461
beacon, radio-range, 443–449
Bellini-Tosi, 441
Beverage, 465–468
counterpoise, 433
coupling circuits, 90
coupling methods to transmission lines, 476–480, 487, 488
current and voltage distribution in, 416–418
Antennas, "current-feed," 477
directional, 435–473
dummy, for receiver tests, 407
effective height of, 81, 438
field distribution of, 420–429
Hertz, 475
horizontal rhombic, 471–473
image, 418
induction field around, 420
inverted-V, 468–471
long-wire, directional, 463–468
loop, 436–439
for transmitting, 443–449
multiple-tuned, 473, 474
radiation field around, 420
radiation resistance, 429–433, 475, 476
for radio direction finders, 439–443
radio-range beacons, 443–449
errors in, 442–443
reflectors for, 436, 452–457
standard, for receiver tests, 407
tilted-wire, 468–471
transmission lines for, 474–491
"voltage-feed," 477
wave, 465–468
"Zeppelin," 477
Anti-resonance, 38
Arc, Poulsen, 306
Array, antenna, 429
gain of, 462
polar diagrams of, 454–457
practical aspects of, 461–463
principles of, 449–454
types of, 458–461
Artificial antenna, for receiver tests, 407
Attenuation constant, 482
Audio-frequency amplifiers (*see* Amplifiers)
Audio-frequency transformers, 163–169
characteristics of, 167
equivalent circuit of, 165
Audion, 116, 344
Autodyne reception, 372
Automatic volume control, 394–398

Autotransformer, coupling for modulator tube, 324
  in oscillator tank, 264
Average value, 3

B

Balance, conditions for, in neutralizing circuits, 255–258
Band-pass filters, 86
Bar, definition of, 171
Barkhausen oscillations, 309–311
Beacons, radio-range, 443–449
Beat frequency, 370, 371
Bellini-Tosi antenna, 441
Bending of electromagnetic waves, 491–495
Beverage antenna, 465–468
Bias, grid, 137
Blocking, in oscillators, 279
  in regenerative detectors, 374
Blocking condensers, 154, 161, 322, 361
Bombardment, positive-ion, 127–129
Bridge neutralizing circuits, 253–258, 293, 294
  adjustments of, 258, 293
Broadside arrays, 458, 461
By-pass condenser, 193, 194, 216, 323, 365

C

C battery, 137
C bias, 137
  detector, 360–365
  methods of obtaining, 192–195
C.W., 314
  reception of, 317–373, 382
Capacitance, calculation of, 69
  coupling, 92
  distributed, 49
  grid-input, 154, 230–235
    variation of, 234
  interelectrode, 154, 228, 234
Carrier suppression, 331–336
Cathodes, 121–127
  equipotential, 121
Cathodes, heat-shielded, 122
  heater-type, 121
  oxide-coated, 122
  thoriated tungsten, 123
    carbonized, 124
Characteristic, dynamic, 150, 153
Characteristic, I_p-E_g and I_p-E_p, 130
  curves, dynatron, 308
    pentode, 206
    tetrode, 203
    triode, 131, 132, 152, 275
  static, 150
Characteristic impedance, 465
  definition of, 481
  of transmission line, 478, 484
Charge of condenser through *R* and *L* in series, 96
Choke coils, 49
  iron-core, 67
Coefficient of coupling, 77
  capacitive, 92
Coil antennas, 436–439
  effective height of, 438
    (*See also* Loop antennas)
Coils, bank-wound, 51
  distributed capacitance of, 51
  honeycomb, 51
  inductance of, air core, 48, 49, 53
  iron-core, 55–60, 65–69
  multilayer, 49
  parallel resonant effects in, 52
  Q of, 30, 56
  resistance of, 30
Colpitts oscillator circuit, 268
Compandor, 400–402
Complex quantities, 10–16
Concentric line, 475
  characteristic impedance of, 478, 484
Condensers, 69–74
  aging of, 70
  blocking, 154, 161, 322, 361
  by-pass, 193, 194, 216, 323, 365
  electrolytic, 73
    dry-type, 74
  figure of merit of, 72
  losses in, 72
  straight-line, capacity, 71

Condensers, straight-line, frequency, 72
 wave length, 71
 variable, 71
Conductance, grid, 228–235
 negative, 233, 234
Continuous wave (C.W.), 314
 reception of, 371–373, 382
Conversion transconductance, 390
Converters, pentagrid, 388–390
Counterpoise, 433
Coupled circuits, 77–93
 capacitive coupling, 92
 combinations of inductive and capacitive coupling, 93
  in radio-frequency amplifiers, 246–249
 conditions for maximum secondary current, 83–86
 critical coupling, 85
 curves of secondary current, 87
 free oscillations in, 103
 frequencies of oscillations in, 105
 resonant relations, 82
 sharpness of tuning in, 88
 tuned, in radio-frequency amplifiers, 249–251
Coupled oscillatory circuits, mechanical analogue of, 106
Coupled resonant circuits, 80
Coupling, coefficient of, 77
 capacitive, 92
 critical, 85
 electron, 388
 of transmission line to antenna, 476–480, 487, 488
"Cross-modulation," 341, 390–392
"Cross-talk" (*see* "Cross-modulation")
Crystal, quartz, 294–298
  equivalent electrical circuit, 297
  mounting of, 297
  temperature coefficient, 296
  X- and *Y*-cuts, 295, 296
Crystal detectors, 344
Current, effective value of, 3
 grid, effect on distortion, 147
Current, plate, expressions for, 134–136, 337–340
 thermionic, 118
Cut-off voltage, 129, 134
 remote, tubes having, 392

D

Damped waves, 106–109
Decibel, 170
Demodulation, 314
  (*See also* Detectors)
Detection coefficient, 346
Detectors, 109, 344–378
 autodyne, 372
 crystal, 344
 diode, 365–370
  as automatic volume control, 394–397
  distortion in, 370
  full-wave, 367
  rectification diagram of, 369
 grid-bias, 360–365
 grid-current, 349–355
  distortion in, 355–358
 grid-leak, 349–355
  distortion in, 355–358
 grid-leak power, 358–360
 heterodyne, 370–372
 linear, 345
 oscillating, 372
 plate, 345–349
 regenerative, 372–374
  blocking in, 374
 square-law, 345
 transrectification diagram of, 363, 364
Dielectric constant, 69
Dielectric losses, 70
Direction finders, 439–443
 errors in bearings of, 442, 443
Directional antennas, 435–473
Diode, automatic volume control, 394–397
 detector, 365–370
 plate current of, 119
Direct-current amplifier, 209–211

Discharge of condenser through R and L in series, 102
Distorted waves, 21–25
  effective value of, 23
Distortion, in amplifiers, 146–150
  amplitude, 146
  in detectors, 355–358
  frequency, 147
  graphical determination of, 182, 187, 208
  phase-shift, 147
  types of, 146–150
Distributed capacitance, 49
Double-grid tubes, 198
Duplex tubes, 200
Dust cores, 56, 383
Dynamic characteristic, 150, 153
Dynamic loud-speakers, 411
Dynatron oscillators, 304–306

E

Echo signals, 496, 497
Eddy currents in iron core, 56
Edison effect, 116
Effective height, of antenna, 81
  of loop antenna, 438
Effective value, 3
Electrical images, 418
Electrical units, 415
Electrolytic condensers, 73
  dry-type, 74
Electromagnetic radiation, 419–429
Electromagnetic waves, 413–415
  echo signals, 496, 497
  fading, 497–499
    of broadcast signals, 424, 434
  long-wave transmission, 499
  production of, 413–415
  propagation of, 491, 492
  refraction of, 491–495
  short-wave transmission, 499–502
  skip-distance effect, 495, 496
Electron, charge and mass of, 117
  coupling, 388
  emission of, 117–120
  oscillations of, 309–311
Electron affinity, 118
  of tungsten, 119
Emission of electrons, 117
  photoelectric, 120
  secondary, 120, 279
Emission efficiency, 125
Emitters, comparison of, 125–127
End-fire array, 458
Equivalent circuit of triode, 137, 228
Equivalent grid voltage, 135
Equivalent plate voltage, 134
Equivalent tube circuit, 137
Excitation, grid, 262, 272, 274, 287
  impulse, 108, 111

F

Fading, 497–499
  of broadcast signals, 424, 434
  selective, 498
Feed-back, 194, 212–220
  acoustic, 213
  in audio-frequency amplifiers, 212–220
  caused by common source of plate voltage, 214–216
  caused by self-bias, 194
  in detectors, 372–374
  "motor-boating," 215
  prevention of, 215–217, 253–258
  in radio-frequency amplifiers, 253
  reversed, 218–220
Feeder (*see* Transmission lines)
Fidelity, of radio-frequency amplifiers, 244–246
  tests on, in receiving sets, 408–410
    (*See also* Distortion)
Field, induction, 420
  radiation, 420
Field distribution, of quarter-wave antenna, 420–424
  of vertical antenna, 424–429
Filament (*see* Cathodes)
Filters, in audio amplifiers, 216
  for B-supply in receivers, 406, 407
  band-pass, 86
  in detectors, 357
  low-pass, 357
  prevention of feed-back by, 215, 216
Fleming valve, 116, 344

Fourier series, 21
  oscillator plate current expressed by, 277
Frequency converters, superheterodyne, 386–390
    electron-coupled, 388–390
Frequency, of free oscillations, 102
  parallel resonant, 34, 36, 45
  radio and audio, 2
  series resonant, 28
  stability of oscillators, 299–301
Frequency distortion, 147
Frequency modulation, 314, 317
Frequency multipliers, 291–293

G

Gain, expressed in decibels, 170
  translation, 390
Gap, nonsynchronous, 110
  quenched, 108
  rotary, 107
Gas, effects of, in vacuum tubes, 127–129
  ionizing potential of, 127
Gonimometer, 441
Grid, action of, 129, 130
Grid-bias, detectors, 360–365
  voltage, 137
    methods of obtaining, 192–195
Grid conductance, 228–235
  negative, 233, 234
Grid current, effect on distortion, 147
Grid-excitation voltage, 262, 272, 274, 287
Grid-input capacitance, 230–235
  variation of, 234
Grid-leak detectors, 349–355, 358–360
Grid modulation, 326–327
Ground wave, optimum conditions for, 433–435

H

Harmonics, 21–25
  in amplifiers, 146
  in oscillators, 270, 291–293
    (*See also* Distortion)
Hartley oscillator circuit, 266
Hazeltine neutralizing circuit, 256
Heising modulation, 319
Heterodyne detection, 301, 370–372
High-angle radiation, 433, 434
Hum, from alternating filament current, 191
Hybrid transformer, 221
Hypernik, 62
Hysteresis, 59
  distortion caused by, 59
  loops, 64
    displaced, 66

I

I.C.W., 315
Ideal transformer, 175, 176
Ignition interference, 111
  suppressors for, 111
Image-frequency interference, 384, 385
Images, electrical, 418
Impedance, 6
  characteristic, 465
    definition of, 481
    of transmission line, 478, 484
  definition of, 6
  input, of a triode, 228–235
  matching, 175, 176
    with quarter-wave-length line, 486–488
  reflected, 80
  surge, 102, 465
Impulse excitation, 108, 111
Incremental permeability, 65
  effect of air gap on, 67
  effect of d.-c. saturation, 67
Inductance, calculation of, in coils, 48, 49, 53
    series aiding or opposing, 78
  definition of, 48
  of iron-core coils, 55–60, 65–69
    with d.-c. and a.-c. excitation, 67
  of multilayer coils, 53
  mutual, 77
  of single-layer solenoid, 49
Induction field, 420

Inductive reactance, 6
Input impedance, grid, 228–235
  of a transformer, 80
Interchannel-noise suppression, 398–400
Interelectrode capacitance, 154, 228, 243
Interference, from "cross-modulation," 341, 390–392
  from electrical sparks, 110–111
  from ignition systems, 111
    suppressors for, 111
  image-frequency, 384, 385
  from static, 111–112
Intermediate frequency, 371
  in superheterodynes, 381, 385, 386
Inverted speech, 335
Ionization by collision, 127
Ionizing potential, 127
Ionosphere, 492–495
  layers composing, 493
  reflection from, 494
Ions, mass of, 127
  production of, 127
Iron-core coils, 55–60, 65–69

K

Kennelly-Heaviside layer, 491
  reflection from, 494
  virtual height of, 493
Konel metal, 123

L

Layer, Kennelly-Heaviside, 491
Layers, in ionosphere, 493
Lecher wires, 310
Line (*see* Transmission lines)
Linear detectors, 345
Litzendraht wire (litz), 55
Loop antennas, 436–439
  as direction finders, 439–443
  effective height of, 438
  radio-range beacons, 443–449
Losses, dielectric, 70
  in magnetic cores, 56, 60, 61, 64
Loud-speakers, 410–412
  balanced-armature, 410
  dynamic, 411

M

Magnetic alloys, 60–64
Magnetic modulator, 318
Magnetostriction oscillators, 299
Magnetron, 307–309
  split-anode, 309
Mass, of electrons, 117
  of ions, 127
Master oscillator, 268
Maximum power, considerations for, 172–177
Maximum undistorted output, 180, 181
Meissner oscillator circuit, 268
Mercury-vapor rectifiers, hot-cathode, 128
Microphones, "zero-level" in, 171
Microphonic effects in amplifiers, 170, 213
"Mixer" tube, superheterodyne, 386
Modulated amplifiers, 321–337
  Class C, 290, 321–323
Modulated oscillators, 319–321
  equivalent circuit of, 319
Modulation, 313–342
  absorption, 318
  amplitude, 315–317
    methods of obtaining, 318
  balanced, 331–334
  Class C amplifiers, 321–323
  energy relations, 317, 320
  factor, 315, 330
  frequency, 314, 317
  grid, 326, 327
  grid-current, 333
  Heising, 319
  by nonlinear impedances, 327–330
  phase, 314, 317
  plate, 319–323
    of Class C amplifiers, 321–323
    coupling circuits for, 323–326
    equivalent circuit of, 319
  suppressed-carrier, 331–333
  suppressor-grid, 336
  types of, 313–315
  van der Bijl, 330

Modulators, balanced, 331–334
  coupling circuits for, 323–326
  magnetic, 318
  tubes suited for, 319
"Motor-boating" in amplifiers, 215
Moving-coil loud-speaker, 411
Multilayer coils, 49
Multiple signals, 496, 497
Multiple-tuned antennas, 473, 474
Multipliers, frequency, 291–293
Multivibrator, 303, 304
Mutual conductance, 132
  measurement of, 141
Mutual inductance, definition of, 77
  in tuned radio-frequency amplifiers, 243–246

N

Negative resistance in triodes, 233, 234
Neutralizing circuits, 253–258, 293, 294
  adjustments of, 258, 293
  for push-pull amplifiers, 294
Neutrodyne circuit, 256
Noise, radio, from electrical sparks, 110–111
  from static, 111–112
Noise suppression, interchannel, 398–400
Nonresonant transmission lines, 477–481
  characteristic impedance of, 478, 484
  termination of, 478, 479
Nonsinusoidal waves, 21–25
Notation, vacuum tube, 136

O

Open-wire line, 475
  characteristic impedance of, 478, 484
Optimum coupling, 82
  in tuned radio-frequency transformers, 242
Oscillations, approximate conditions for, in triodes, 265
Oscillations, in coupled circuits, 103
  electron, 309, 310
  free, criterion for, 100
  parasitic, 268, 269
  in tuned amplifiers, 253–258
Oscillators, 261–283, 298–311
  adjustments, 263, 264, 272
  Barkhausen, 309–311
  beat-frequency, 301–303
  blocking of, 279
  circuit calculations, 273–278
  crystal, 294–298
    (*See also* Crystal, quartz)
  current and voltage relations, 269–273
  dynatron, 304–306
  electron, 309–311
  frequency stability of, 299–301
  magnetostriction, 299
  magnetron, 307–309
    split-anode, 309
  plate efficiency of, 281
  plate-modulated, 319–321
  Poulsen arc, 306
  power relations, 278–283
  push-pull, 269
  secondary emission in, 279
  starting of, 263
  typical circuits, 266–268
  unstable operation in, 272, 279
Output, maximum undistorted, 180, 181
Oxide-coated emitters, 122, 123

P

Parabolic reflectors, 436
Parasitic oscillations, 268, 269
Parasitic reflecting antenna, 457
Parallel circuits, 9
Parallel resonance, 32
  conditions for maximum impedance, 37, 42
  conditions for unity power factor, 36, 43, 45
  effect of resistance, 35–44
  sharpness of, 44
  vector diagrams of, 42
Pentagrid converters, 388–390

Pentodes, 204–209
distortion in, 208
as modulated amplifiers, 336
as radio-frequency amplifiers, 242, 243, 381
Performance tests on receiving sets, 407–410
Permalloy, 60
Perminvar, 63
Permeability, 56
incremental, 65
effect of air gap on, 67
effect of d.-c. saturation on, 67
of magnetic alloys, 60–64
values of, 57
Phase, shift distortion, 147
Phase modulation, 314, 317
Phase velocity, 468, 491, 492
Photocells, 121
Photoelectric emission, 120
Piezoelectric, 294–298
(*See also* Crystal, quartz)
Plate current, expressions for, 134–136, 337–340
Plate detection, 345–349
Plate resistance, 133
measurement of, 141
Polar vectors, 16
Positive-ion bombardment, 127–129
Poulsen arc, 306
Power, alternating current, 19
apparent, 20
Power amplifiers, Class A, 177–179, 184–190, 204–209
Class AB, 200, 201
Class B, 195–200, 287, 288
conditions for maximum output, 180, 190
pentodes used as, 204–209
push-pull, 184–190, 195–201
radio-frequency, 269–294
approximate characteristics, 283–285
circuit calculations, 273–278
Class B, 287, 288
Class C, 289–291
current and voltage relations in, 269–273
Power amplifiers, radio-frequency, modulated, 321–323
neutralization of, 293, 294
plate efficiency of, 281
power relations, 278–283
secondary emission in, 279
tuning adjustments, 285, 286
(*See also* Amplifiers)
Power factor, 19
Propagation constant, 482, 485
Push-pull amplifiers, 184–192, 195–201
modulators, 325
oscillators, 269

Q

Q, of coils, 30, 56
Quarter-wave antenna, 417
field distribution of, 420–424
transmission lines, 486–488
Quartz crystals, 294–298
(*See also* Crystal, quartz)
Quenched gap, 108

R

Radiation, electromagnetic, 419–429
field of, 420
high-angle, 433, 434
resistance, 429–433, 475, 476
(*See also* Electromagnetic waves)
Radio compass, 439–443
errors in bearings of, 442, 443
Radio-frequency amplifiers (*see* Amplifiers)
Radio interference, from electrical sparks, 110
from static, 112
Radio-range beacons, 443–449
Radio waves (*see* Electromagnetic waves)
Reactance, 6
apparent, 36, 44
condensive, 7
inductive, 6
Reactivation of thoriated filament, 124

Receivers, all-wave, 407
regenerative, 373, 380
standard output for, 407
superheterodyne, 381–390, 404–407
circuit diagram of, 405
frequency converters for, 386–390
image-frequency interference in, 381–385
intermediate frequency, 381, 385, 386
superregenerative, 374, 375
tuned radio-frequency, 380, 381
types of, 380
Receiving sets, performance tests, 407–410
Receiving systems, 380–410
Rectifier, mercury-vapor, 128
for receiving sets, 407
Tungar, 127
(*See also* Detectors)
Reflected impedance, 80
from antenna, 90
Reflecting transformer, 467
Reflectors, for antenna arrays, 452, 457
parabolic, 436
parasitic, 457
Regeneration, 212
acoustic, 213
in audio-frequency amplifiers, 212–220
caused by common source of plate voltage, 214–216
caused by self-bias, 194
in detectors, 372–374
"motor-boating," 215
in radio-frequency amplifiers, 253
Regenerative amplification, 212, 213, 217–220
detectors, 372–374
Resistance, due to skin effect, 54
negative, in triodes, 233, 234
of parallel resonant circuit, 43
plate, 133
radiation, 429–433, 474–476
of single-layer coil, 30
Resistance, of variable condensers, 72
Resonance, 9
parallel, 32, 44
series, 28
sharpness of, 29, 44
vector diagrams of, 19, 42
Resonant transmission lines, 475–477
(*See also* Transmission lines)
Reversed feed-back, 194, 218–220
Rhombic antennas, horizontal, 471–473
Rice neutralizing circuit, 255, 293, 294
Root-mean-square value, 3
Rotary gap, 107
nonsynchronous, 110

S

Screen-grid tubes, 201–204
(*See also* Tetrodes)
Secondary emission, 120, 279, 304
in tetrodes, 202, 305
Selectivity of tuned amplifiers, 243–245, 252
Self-bias, in amplifiers, 192–195
Series circuits, 4–8
Series resonance, 28
sharpness of, 29
vector diagram, 19
Sharpness of resonance, 29, 44
in coupled circuits, 88
Side bands, 316
energy contained in, 317
Signal generator for receiver tests, 408
Sine wave of voltage, 1
average value of, 3
effective value of, 3
Single-side-band transmission, 334–336
Skin effect, 53
Skip-distance effect, 495–496
Sky wave, radio-compass errors caused by, 442
Space charge, 119–122

Space charge, neutralization of, by positive ions, 128
Space-charge grid, 204
Spark gap, types of, 107, 108
Spark transmitters, 106
limitations of, 111
power requirements of, 110
Speech, inverted, 335
Standard antenna, for receiver tests, 407
Standing waves, 416
Static, radio interference caused by, 111–112
Static characteristic, 150
Superheterodyne, 371, 381–390, 404–407
Superregeneration, 374, 375
Suppressed carrier, 331–336
Suppressor grid, 205
Surge impedance, 102
(*See also* Characteristic impedance)

T

Taylor's series, 135, 136
plate current expressed by, 337–340
Telephone repeaters, 220–222
Tests, on receiving sets, 407–410
Tetrodes, 201–204
as Class A amplifier, 199, 201–204
as Class B amplifier, 198
characteristics of, 203
as dynatron oscillator, 305
secondary emission in, 202
with space-charge grid, 204
Thermionic current, 118
effects of gas on, 127
saturation value, 118
Thermionic emission, 118
Thermionic tube, history of, 115–117
Thermionic work function, 118
Thoriated filaments, 123–125
carbonized, 124
loss of emission at low temperatures, 126
reactivation of, 124
Tone control, 403, 404
"Tracking frequencies," 388
Transconductance, 132
conversion, 390
Transformers, audio-frequency, 166–169
hybrid, 221
ideal, 175, 176
impedance-matching, 175
input impedance of, 80
intermediate-frequency, 383, 384
reflecting, 467
tuned radio-frequency, 239–251, 381
untuned, 238
Transient oscillations, in coupled circuits, 103
R, *L*, and *C* in series, 96–103
Translation gain, 390
Transmission, single-side-band, 334–336
spark, 106–110
Transmission lines, 471–491
characteristic impedance of, 478. 484
concentric, 475
coupling methods to antenna. 476–480, 487, 488
measurements on, 489
nonresonant, 477–481
open-wire, 475
quarter-wave, 486–488
as impedance-matching device. 487, 488
for reception, 489–491
resonant, 475–477
single-wire, 479, 480
termination of, 478, 479
theory of, 481–486
types of, 474, 475
Transrectification diagram of detectors, 363, 364
Triode, equivalent circuit of, 137, 228
Triode constants, measurement of, 140–142
from structural dimensions, 139
Triodes, 129
characteristic curves of, 131, 132, 152, 275

Triodes, constants of, 131–134
Triple-grid tubes, 205
Tubes, diode, 119, 365–370, 394–397
  double-grid, 198
  duplex, 200, 367
  fundamental properties of, 115–142
  hot-cathode mercury-vapor, 122, 128
  "mixer," superheterodyne, 386
  pentagrid converter, 388, 390
  pentodes, 204–209
  "supercontrol," 390–394
  tetrode, 201–204
  triple-grid, 205
  variable-mu, 390–394
  water-cooled, 117, 282
Tuned amplifiers, 239–252
  effect of mutual inductance, 243–246
  inductive and capacitive coupling in, 246–249
  neutralization of, 253–258
  optimum coupling in, 242
  parallel resonant, 239
  selectivity of, 245
  transformer-coupled, 239–243
  using tuned-coupled circuits, 249–251
    (*See also* Amplifiers)
Tungar rectifier, 127
Tungsten, thoriated 123–125
Tuning adjustments, power amplifiers, 285, 286

U

Undistorted output, maximum, of amplifiers, 180, 181
Units, electrical, 415
Unstable operation of oscillators, 272, 279

V

Vacuum-tube notation, 136
Vacuum-tube voltmeters, 375–378
Vacuum-tubes, fundamental properties of, 115–142
Variable condensers, 71
  resistance of, 72
  types of, 71–72
van der Bijl modulator, 330
Variable-mu tubes, 390–394
Vector operations, 10–16
Vectors, polar, 16
Velocity, group, 491, 492
  phase, 468, 491, 492
Vertical antenna, field distribution of, 424–429
Voltage amplification, 144
  expressed in decibels, 170
  graphical determination of, 152
  impedance-coupled, 157–161
  measurement of, 171
  resistance-coupled, 154–157
  with resistance load, 151
  transformer-coupled, 163–169
  tuned radio-frequency, 241–243, 250, 251
Voltmeters, vacuum-tube, 375–378
Volume control, acoustically compensated, 402, 403
  automatic, 394–398
  compandor, 400–402

W

Water-cooled tubes, 117, 282
Wave, continuous, (C.W.), 314
  interrupted continuous (I.C.W.), 315
Wave antenna, 465–468
Wave length, 113, 416
  choice of, for transmission purposes, 501, 502
  critical, 499
  of transmission lines, 416, 475–477
Wave meter, 113
Waves, damped, 106–109
  distorted, 21–25
  electromagnetic, 413–415
    (*See also* Electromagnetic waves)

Z

"Zero level," in microphones 171